

STIC-Biotech/ChemLib

89/22

From: Slobodyansky, Elizabeth  
Sent: Friday, March 14, 2003 3:56 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/744,125

Please search for case 09/744,125:

*EL*

SEQ ID NOS: 1 and 2 against commercial and interference databases.

Thank you.

*Elizabeth Slobodyansky, PhD*

Primary Examiner

Art Unit 1652  
CM1 10D11  
703-306-3222

mail box 10D01

Mary Jane Ruhl  
Tech. Info. Specialist, STIC  
TC-1600  
CM-1, Room 6A-06  
Phone: 605-1155

RECEIVED  
MAR 14 2003  
STIC

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 3/8  
Date Completed: 3/31  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**This Page Blank (uspto)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 27, 2003, 05:22:50 ; Search time 446 Seconds  
(without alignment)  
7841.600 Million cell updates/sec

Title: US-09-744-125A-1

Perfect score: 1553  
Sequence: 1 gaattcgccagcagcgtcg999.....aaaaaaaaaacctcgag 1553

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553	100.0	1553	21	AAZ58953 Human MIF1 protein
2	1535.8	98.9	1914	21	AAZ58958 Human MIF1 protein
3	414.4	26.7	10579	22	ABA20014 Human nervous syst
4	414.4	26.7	10579	22	AAK70045 Human immune/haema
5	414.4	26.7	28995	22	ABA20015 Human nervous syst
6	414.4	26.7	28995	22	AAK70046 Human immune/haema
7	414.4	26.7	28995	22	AAK79967 Human immune/haema
8	414.4	26.7	28995	22	AAK85213 Human immune/haema
9	351.4	22.6	1905	23	ABL02115 Drosophila melanog

C	10	285.6	18.4	4027	23	ABL02114
	11	159.4	10.3	8204	22	AAK79966
	12	159.4	10.3	8204	22	AAK85212
	13	121	7.8	144	16	AA723953
	14	71.6	4.6	457	22	AAK87783
	15	71.6	4.6	457	22	AAK81700
	16	71.6	4.6	457	22	ABN90055
	17	60.6	3.9	20974	23	AAK92595
	18	60.2	3.9	65	24	ABN52890
	19	60	3.9	60	24	ABN59396
	20	58	3.7	2108	24	ABU67774
	21	50.2	3.2	60	24	ABN59283
	22	50.2	3.2	60	24	ABN59397
	23	49.4	3.2	51	22	AAJ30900
	24	49.4	3.2	799	19	AAV55831
	25	49.4	3.2	1925	20	AAK90924
	26	49.4	3.2	1926	21	AAK50254
	27	49.4	3.2	1926	22	AAK82902
	28	49.4	3.2	2580	21	AAK75454
	29	49.4	3.2	2580	24	AAK64275
	30	49.4	3.2	5452	20	AAK90923
	31	49.4	3.2	8705	20	AAK23378
	32	49.4	3.2	9600	19	AAK21683
	33	49.4	3.2	10380	20	AAK22248
	34	49.4	3.2	10596	14	AAK051731
	35	49.4	3.2	10596	17	AAK40348
	36	49.4	3.2	10596	20	AAK15650
	37	49.4	3.2	16080	21	AAK59553
	38	48.6	3.1	3489	21	AAK10290
	39	48.6	3.1	3489	22	AAK82901
	40	48.6	3.1	3489	24	AAK93487
	41	48.6	3.1	32207	24	AAK73805
	42	48.6	3.1	137507	19	AAV19941
	43	47.6	3.1	16442	18	AAK83006
	44	47.2	3.0	6522	24	ABU33051
	45	47	3.0	5120	22	AAK84677

#### ALIGNMENTS

RESULT 1	AAZ58953	AAZ58953 standard; cDNA, 1553 BP.
XX	AAZ58953	
AC	AAZ58953	
XX	AAZ58953	
DT	08-MAY-2000	(first entry)
XX	08-MAY-2000	
DE	Human MIF1 protein encoding cDNA (plasmid pCM480).	
XX	MEK kinase; MEK interacting forkhead associated protein; MIF1, MEK;	
KW	FHA protein; forkhead associated protein; tumour; angiogenesis; human;	
KW	psoriasis; cancer; anti-inflammatory; anti-asthmatic; immunosuppressant;	
KW	vasotropic; neuroprotective; antiarthritic; antiviral; ss.	
OS	Homo sapiens.	
XX		
FX	Key	Location/Qualifiers
FT	mat_peptide	2..1174
XX		/tag= a
XX		
XX	WO200005362-A1.	
XX		
XX	03-FEB-2000.	
PD		
XX	21-JUL-1999;	99WO-EP05142.
PF		
XX	21-JUL-1998;	98US-0093590.
PR		
XX	(RHON ) RHONE-POULENC RORER SA.	
PA		
XX		
XX	Marcireau C, Multon M, Polard-Houeset V;	
PI		

Drosophila melanog  
Human immune/haema  
Human immune/haema  
Human gene signatu  
Human digestive sy  
Human liver asocci  
Human liver antige  
DNA encoding novel  
Mouse spliced tran  
Human spliced tran  
Oesophagus cancer  
Human spliced tran  
Human spliced tran  
Human SNP oligonuc  
Nucleotide sequenc  
Epstein-Barr virus  
Epstein Barr virus  
EBV tethering prot  
Nucleotide sequenc  
Epstein-Barr virus  
Anti-sense strand  
Vector pshuttle DN  
Vector plasmid pCM  
Nucleotide sequenc  
Plasmid pCisEBON f  
Plasmid pCisEBON f  
Nucleotide sequenc  
DNA clone pCK Cl.  
Kaposi's sarcoma-a  
Nucleotide sequenc  
Kaposi's sarcoma-a  
KSHV LUR DNA (nucl  
KSHV long unique c  
Partial mouse WRN  
Human immune syste  
DNA sequence of hu

XX MPI: 2000-195102/17.  
 DR P-PSDB; AAY77554.  
 XX  
 PT New MEK kinase interacting forkhead associated protein (MIF1) useful to  
 PT treat or diagnose, e.g. inflammation and tumors, and to identify its  
 PT specific modulators, to regulate MEK kinase activity -  
 XX  
 PS Claim 3; Page 64-67; 78pp; English.  
 XX  
 CC The invention provides MEK kinase (MEKK) interacting forkhead associated  
 CC (FHA) protein (MIF1). MIF1 is useful for screening for specific  
 CC modulators (potential therapeutic agents) and to reduce MEK activity in  
 CC cells. Antibodies specific to MIF1 are useful as diagnostic immunoassay  
 CC reagents to detect expression of MIF1, for purification of MIF1 and as  
 CC therapeutic (ant)agonists. The MIF1 nucleic acids are useful for:  
 CC recombinant production of MIF1, either in cultured cells or in vivo (gene  
 CC therapy); as source of probes and primers for detecting or quantifying  
 CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source  
 CC of therapeutic antisense sequences (used to increase MEK activity in  
 CC cells); and to identify inhibitors of MIF1. Regulation of MEK activity  
 CC via MIF1 regulation is useful for treatment of inflammation, asthma,  
 CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic  
 CC syndrome, neurodegeneration, tumours, angiogenesis, rheumatoid arthritis,  
 CC psoriasis or persistent viral infections. The present sequence represents  
 CC a cDNA (plasmid pCM480) encoding a MIF1 protein.

XX Sequence 1553 BP; 367 A; 457 C; 423 G; 306 T; 0 other;

Query Match 100.0%; Score 1553; DB 21; Length 1553;

Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGACGACGAGTGGGGGTGACACAGGGCGCTGTTGGGGAGTGAACCTCTCCAG 60  
 DB 1 GAATTCGGACGACGAGTGGGGGTGACACAGGGCGCTGTTGGGGAGTGAACCTCTCCAG 60  
 QY 61 TGAAGAAAAGAAAGTATCCAAAGCCCCCAGACACTCTGTGCCACCCAGCCAGCCCAAC 120  
 DB 61 TGAAGAAAAGAAAGTATCCAAAGCCCCCAGACACTCTGTGCCACCCAGCCCAAC 120  
 QY 121 CCTGGACTCACCAAGCGTGTGAAGAAAGTAAACAGCCACTTCAGGTGACCAAGATCT 180  
 DB 121 CCTGGACTCACCAAGCGTGTGAAGAAAGTAAACAGCCACTTCAGGTGACCAAGATCT 180  
 QY 121 CCTGGACTCACCAAGCGTGTGAAGAAAGTAAACAGCCACTTCAGGTGACCAAGATCT 180  
 DB 121 CCTGGACTCACCAAGCGTGTGAAGAAAGTAAACAGCCACTTCAGGTGACCAAGATCT 180  
 QY 181 GGGCGCGTGAAGCGTGAATGACCTCTGCTCATTAATGCTGTGTGCAACCAACA 240  
 DB 181 GGGCGCGTGAAGCGTGAATGACCTCTGCTCATTAATGCTGTGTGCAACCAACA 240  
 QY 241 CCTGACCTCCGTCACCTGTGGCGGTGAATTCAGCTGCCCTTCACCTTGGAGGTGCA 300  
 DB 241 CCTGACCTCCGTCACCTGTGGCGGTGAATTCAGCTGCCCTTCACCTTGGAGGTGCA 300  
 QY 301 GGAGCGTTGGTACGCGCTGCTCTAGATCCTGTCATCTCAAGTTGGCTGTCAAGCCAT 360  
 DB 301 GGAGCGTTGGTACGCGCTGCTCTAGATCCTGTCATCTCAAGTTGGCTGTCAAGCCAT 360  
 QY 361 GAGGAGCTGCAACCAAGAGCTATTGACAGCCATCCAGAGCAAGGCCCTGTTAGCAAGGC 420  
 DB 361 GAGGAGCTGCAACCAAGAGCTATTGACAGCCATCCAGAGCAAGGCCCTGTTAGCAAGGC 420  
 QY 421 TGAAGAGAGCTGCTGAGCAAAAGTGGATGACACAGCCAGCCCACTTTGAGACCTTCCA 480  
 DB 421 TGAAGAGAGCTGCTGAGCAAAAGTGGATGACACAGCCAGCCCACTTTGAGACCTTCCA 480  
 QY 481 GGAACCTGTCAGACACACCTGATGCTTCTACCTGAGCCGTAACCGCAAGCCCTTGA 540  
 DB 481 GGAACCTGTCAGACACACCTGATGCTTCTACCTGAGCCGTAACCGCAAGCCCTTGA 540  
 QY 541 GGCCCACTGGCAGCTCATGAAGCAGTATTAACCTGTCGAGGACCAAGAGTGCAGCCGT 600  
 DB 541 GGCCCACTGGCAGCTCATGAAGCAGTATTAACCTGTCGAGGACCAAGAGTGCAGCCGT 600

QY 601 GCCCAAGGGGACCAAGTGTCTAACTTCTGTGATGACAGAGCACTGATGATGACAGTAA 660  
 DB 601 GCCCAAGGGGACCAAGTGTCTAACTTCTGTGATGACAGAGCACTGATGATGACAGTAA 660  
 QY 661 GCTCAAGGACATGCGAGATGAGGTCTTGGAAACATGAGCTGATGTGCTGACCGGCCCA 720  
 DB 661 GCTCAAGGACATGCGAGATGAGGTCTTGGAAACATGAGCTGATGTGCTGACCGGCCCA 720  
 QY 721 GAAGGAGAGATTCGGGACGCTGGGACAGAACTGATAAAGTGGGAGGTCTAGTGAACAG 780  
 DB 721 GAAGGAGAGATTCGGGACGCTGGGACAGAACTGATAAAGTGGGAGGTCTAGTGAACAG 780  
 QY 781 CATCAAGGACATGAGCTCTCCGACTTCGACCAACAGCACTAGCGAGTGTGCGGGCCG 840  
 DB 781 CATCAAGGACATGAGCTCTCCGACTTCGACCAACAGCACTAGCGAGTGTGCGGGCCG 840  
 QY 841 CATGTCGCGTACCTGATGCGTGTGAGATCAACCTGCGGACAGCAACAGAGATTA 900  
 DB 841 CATGTCGCGTACCTGATGCGTGTGAGATCAACCTGCGGACAGCAACAGAGATTA 900  
 QY 901 CCAAGTTGATGTGACCTGTCTTGGAGGTCGCCGCTGGAAAGATATCCGGAAACAGG 960  
 DB 901 CCAAGTTGATGTGACCTGTCTTGGAGGTCGCCGCTGGAAAGATATCCGGAAACAGG 960  
 QY 961 TGTCAATCAAGCTGAAGAACAAACGATGATTTCTTCAATTCGCAATGAGGTGCAAGCCCAT 1020  
 DB 961 TGTCAATCAAGCTGAAGAACAAACGATGATTTCTTCAATTCGCAATGAGGTGCAAGCCCAT 1020  
 QY 1021 CTACATCGATGACGCGCGGTGCTGTGTGCTCCAAATGCGCTCAGCAACACTCTGT 1080  
 DB 1021 CTACATCGATGACGCGCGGTGCTGTGTGCTCCAAATGCGCTCAGCAACACTCTGT 1080  
 QY 1081 GGTGAGATTCGACACCTGCTGCTTCTTCTTCAACCAAGACCTCATGCCCCAT 1140  
 DB 1081 GGTGAGATTCGACACCTGCTGCTTCTTCTTCAACCAAGACCTCATGCCCCAT 1140  
 QY 1141 CAGGCTGAGGCTGCGCAAGATCACACACAGTGAAGTGGGAGGACTGTGGGCC 1200  
 DB 1141 CAGGCTGAGGCTGCGCAAGATCACACACAGTGAAGTGGGAGGACTGTGGGCC 1200  
 QY 1201 TCTCGGCGCTGTTTCCCTGCACTCCAGCCCCCTTGAAGCTGGGAACTCAGGCTCTGGA 1260  
 DB 1201 TCTCGGCGCTGTTTCCCTGCACTCCAGCCCCCTTGAAGCTGGGAACTCAGGCTCTGGA 1260  
 QY 1261 AAAAAGCTGGGAGTGGAGGCTCAAGCTGGGGCAATGATTTGAGCGAGGAT 1320  
 DB 1261 AAAAAGCTGGGAGTGGAGGCTCAAGCTGGGGCAATGATTTGAGCGAGGAT 1320  
 QY 1321 AGGCTGGGCTTGTGAAGCCAGAGAGCTGAGAACCTCAGGCTTCCCTAGATCCAGAG 1380  
 DB 1321 AGGCTGGGCTTGTGAAGCCAGAGAGCTGAGAACCTCAGGCTTCCCTAGATCCAGAG 1380  
 QY 1381 CCCCTGCCCATCTTCTCTCTAATAAACCAACCTTACCCCCATTCTACCCCATTTGCC 1440  
 DB 1381 CCCCTGCCCATCTTCTCTCTAATAAACCAACCTTACCCCCATTCTACCCCATTTGCC 1440  
 QY 1441 ACCTTCACTCTCTGTGTTCACAGCTGATTAAGCTCAGACTCTTCTTTATGTTTTCTTT 1500  
 DB 1441 ACCTTCACTCTCTGTGTTCACAGCTGATTAAGCTCAGACTCTTCTTTATGTTTTCTTT 1500  
 QY 1501 TGTAAATATAAAGCAACAGGTTCCAAAGTAAATAAATAAATAAATAAATAAATAAATAA 1553  
 DB 1501 TGTAAATATAAAGCAACAGGTTCCAAAGTAAATAAATAAATAAATAAATAAATAAATAA 1553

RESULT 2  
 AA258958  
 ID AA258958 standard; cDNA; 1914 BP.  
 XX AA258958;  
 AC  
 XX  
 DT 08-MAY-2000 (first entry)  
 XX



DE Human MIF1 protein encoding cDNA (plasmid pCM577).  
XX MEK kinase; MEKK interacting forkhead associated protein; MIF1; MEKK;  
XX FHA protein; forkhead associated protein; tumour; angiogenesis; human;  
KM psoriasis; cancer; anti-inflammatory; anti-asthmatic; immunosuppressant;  
KM vasotrophic; neuroprotective; antiarthritic; antiviral; ss.  
XX  
OS Homo sapiens.  
PH Key location/Qualifiers  
FT CDS 147..1535  
FT /\*tag= a  
XX  
XX MO200005362-A1.  
XX  
XX 03-FEB-2000.  
XX  
XX 21-JUL-1999; 99WO-EP05142.  
XX  
XX 21-JUL-1998; 98US-0093590.  
XX  
XX (RHON ) RHONE-POULENC RORER SA.  
XX  
XX Marcireau C, Multon M, Polard-Houssset V;  
XX WPI: 2000-195102/17.  
XX P-PSDB; AAY77555.  
XX  
XX New MEK kinase interacting forkhead associated protein (MIF1) useful to  
PT treat or diagnose, e.g. inflammation and tumours, and to identify its  
PT specific modulators, to regulate MEK kinase activity -  
XX  
XX  
XX Claim 5; Page 70-73; 78pp; English.  
XX  
XX The invention provides MEK kinase (MEKK) interacting forkhead associated  
CC (FHA) protein (MIF1). MIF1 is useful for screening for specific  
CC modulators (potential therapeutic agents) and to reduce MEKK activity in  
CC cells. Antibodies specific to MIF1 are useful as diagnostic immunosay  
CC reagents to detect expression of MIF1, for purification of MIF1 and as  
CC therapeutic (antagonists). The MIF1 nucleic acids are useful for:  
CC recombinant production of MIF1, either in cultured cells or in vivo (gene  
CC therapy); as source of probes and primers for detecting or quantifying  
CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source  
CC of therapeutic antisense sequences (used to increase MEKK activity in  
CC cells); and to identify inhibitors of MIF1. Regulation of MEKK activity  
CC via MIF1 regulation is useful for treatment of inflammation, asthma,  
CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic  
CC syndrome, neurodegeneration, tumours, angiogenesis, rheumatoid arthritis,  
CC psoriasis or persistent viral infections. The present sequence represents  
CC a CDNA (plasmid pCM577) encoding a MIF1 protein.  
XX  
XX  
SQ Sequence 1914 BP; 449 A; 549 C; 534 G; 382 T; 0 other;  
Query Match 98.9%; Score 1535.8; DB 21; Length 1914;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1537; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 15 GTGGGGTGAACACGAGGCGCTGTTGGGGAGTGAACCTCTCTCAAGTGAAGAAAGG 74  
Db 376 GTGGGGTGAACACGAGGCGCTGTTGGGGAGTGAACCTCTCTCAAGTGAAGAAAGG 435  
QY 75 TATCCAAAGCCCCCGACCTCTGTGCAACCGAGCCGACCCCGCTGAGACACCA 134  
Db 436 TATCCAAAGCCCCCGACCTCTGTGCAACCGAGCCGACCCCGCTGAGACACCA 495  
QY 135 AGCGTGTGAAGAAGTAAACGACCTTCAGGTGACCAAGGATCTGGGCGCTGGAAGC 194  
Db 496 AGCGTGTGAAGAAGTAAACGACCTTCAGGTGACCAAGGATCTGGGCGCTGGAAGC 555  
QY 195 CTGCAATGACCTCTCTGCTCATTAATGCTGTGTGCAAGACCAAGACCTGACCTCGTCC 254  
Db 556 CTGCAATGACCTCTCTGCTCATTAATGCTGTGTGCAAGACCAAGACCTGACCTCGTCC 615

QY 255 ACCTGGGCGTGAATTCAGCTGCGCGCTTCACTCTCGGAGGTCCAGAGCGTTGGTAG 314  
Db 616 ACCTGGGCGTGAATTCAGCTGCGCGCTTCACTCTCGGAGGTCCAGAGCGTTGGTAG 675  
QY 315 CCCGCTTACGATTCCTGTCTATCTCCAAAGTTGGGCTGTCAAGCCATGAGGACGCTGCAC 374  
Db 676 CCCGCTTACGATTCCTGTCTATCTCCAAAGTTGGGCTGTCAAGCCATGAGGACGCTGCAC 735  
QY 375 CAGAGGCTATTGACGCCATCCAGAGCAAGGCGCTGTTAGCAAGGCTGAGAGCGAGCTGC 434  
Db 736 CAGAGGCTATTGACGCCATCCAGAGCAAGGCGCTGTTAGCAAGGCTGAGAGCGAGCTGC 795  
QY 435 TGAGCAAGTGGATTCGACCGACCGACCCACTTTGAGAGACCTTCAGAGACCTGTGCA 494  
Db 796 TGAGCAAGTGGATTCGACCGACCGACCCACTTTGAGAGACCTTCAGAGACCTGTGCA 855  
QY 495 GACACCTGATGCTTCTTCACTGTGCGCGGTACCGGAAAGGCGCTGAGGCCACTGCGACG 554  
Db 856 GACACCTGATGCTTCTTCACTGTGCGCGGTACCGGAAAGGCGCTGAGGCCACTGCGACG 915  
QY 555 TCATGAAGCAGTATTTAAGTGTGAGAGCAAGAGTGCAGCGCTGCGCAAGGGGAGCC 614  
Db 916 TCATGAAGCAGTATTTAAGTGTGAGAGCAAGAGTGCAGCGCTGCGCAAGGGGAGCC 975  
QY 615 AAGTGTGAACCTTCTGTGATGACAGAGACCTGATTTGATGACATGATGCTCAAGGACATGC 674  
Db 976 AAGTGTGAACCTTCTGTGATGACAGAGACCTGATTTGATGACATGATGCTCAAGGACATGC 1035  
QY 675 GAGATGAGTCTTGTGAACATGAGCTGATGTGCTGACCGCGCCAGAGCGAGAGATTTC 734  
Db 1036 GAGATGAGTCTTGTGAACATGAGCTGATGTGCTGACCGCGCCAGAGCGAGAGATTTC 1095  
QY 735 GGCAGCTGGAACAGGAAGTGTGATAGTGGCGAGTGTGAGACAGATTCACAGGCATGA 794  
Db 1096 GGCAGCTGGAACAGGAAGTGTGATAGTGGCGAGTGTGAGACAGATTCACAGGCATGA 1155  
QY 795 GCTCTCCGGAATTGACACACAGACACTGCGAGTGTGCGGGCGGACGATGTCGCTGAC 854  
Db 1156 GCTCTCCGGAATTGACACACAGACACTGCGAGTGTGCGGGCGGACGATGTCGCTGAC 1215  
QY 855 TGATGCGCTGCGTGAATCACTCTGGGAGAGCAACCAAGATTAACGATTGATGTCG 914  
Db 1216 TGATGCGCTGCGTGAATCACTCTGGGAGAGCAACCAAGATTAACGATTGATGTCG 1275  
QY 915 ACCGTCTCTGGAAGGTCGCGGCTGGAAGATATCCCGGAAACAAGGTGTCTATCAAGCTGA 974  
Db 1276 ACCGTCTCTGGAAGGTCGCGGCTGGAAGATATCCCGGAAACAAGGTGTCTATCAAGCTGA 1335  
QY 975 AGAACCAAGGTGATTTCTTCAATGCAATGAGGTGCAAGGCGCATCTACATGATGAC 1034  
Db 1336 AGAACCAAGGTGATTTCTTCAATGCAATGAGGTGCAAGGCGCATCTACATGATGAC 1395  
QY 1035 GCGCGGTCTCTGTGCTCCAAATGCGCTCAGCAACCACTCTGTGTGAGATTCGCCA 1094  
Db 1396 GCGCGGTCTCTGTGCTCCAAATGCGCTCAGCAACCACTCTGTGTGAGATTCGCCA 1455  
QY 1095 GCGTGCATTTGCTCTTCTTATCAACAGACCTCATTTGCCCTCATCAGGCGTGAAGCTG 1154  
Db 1456 GCGTGCATTTGCTCTTCTTATCAACAGACCTCATTTGCCCTCATCAGGCGTGAAGCTG 1515  
QY 1155 CCAAGATCACACCAAGTGAAGATGAGGAGGACTGTGAGGCGCTCTCGGCGCTGTTT 1214  
Db 1516 CCAAGATCACACCAAGTGAAGATGAGGAGGACTGTGAGGCGCTCTCGGCGCTGTTT 1575  
QY 1215 CCCGTGCCACTCCAGCGCGCTTGAAGTGGAACTCAAGCTCTTGGAAACCTGCGGAGT 1274  
Db 1576 CCCGTGCCACTCCAGCGCGCTTGAAGTGGAACTCAAGCTCTTGGAAACCTGCGGAGT 1635  
QY 1275 GGGAGGCTCAAGCTCGGGCCATTGATTTGAGACCTTTGAGGAGGATGAGGCTGCGCTTTG 1334  
Db 1636 GGGAGGCTCAAGCTCGGGCCATTGATTTGAGACCTTTGAGGAGGATGAGGCTGCGCTTTG 1695  
QY 1335 TGAAGCCAGCAGAGGCTGAGAACTCAGGCTTCCCTAGATCAAGAGCCCTCCCATCTT 1394

Db 1636 TGAAGCCAGACAGGCTGGAACCTCAGGCTTCCTCCAGTTCAGAGCCCCCTCCCACTCT 1755  
Qy 1395 CCTCTCTCTAAAAACAACCTTACCCCCCATTTACCCCCCATTTGACCTTCACTCTCT 1454  
Db 1756 CCTCTCTCTAAAAACAACCTTACCCCCCATTTACCCCCCATTTGACCTTCACTCTCT 1815  
Qy 1455 GCTCCAGCTGATTTAGCCCACTCTTTTATTTGTTTCTTTGTAATAATAAACG 1514  
Db 1816 GCTCCAGCTGATTTAGCCCACTCTTTTATTTGTTTCTTTGTAATAATAAACG 1875  
Qy 1515 ACCAGGTTCCAAAGTAAAAAATCTCGAG 1553  
Db 1876 ACCAGGTTCCAAAGTAAAAAATCTCGAG 1914  
RESULT 3  
ID ABA20014/C  
AB20014 standard; DNA; 10579 BP.  
XX ABA20014;  
XX  
XX  
XX 23-JAN-2002 (first entry)  
DE Human nervous system related polynucleotide SEQ ID NO 12345.  
XX  
XX Human; nootropic; neuroprotective; cytosstatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vlnnerary;  
KW antiparkinsonian; antischizling; antianaemic; antitachytic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antituber; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO200159063-A2.  
XX  
XX 16-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01334.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225247.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234222.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239933.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.



PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232081.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244671.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251866.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0255678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI, 2001-483426/52.  
  
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
useful for preventing, diagnosing and/or treating cancers and  
metastasis -  
Disclosure: SEQ ID NO 24857; 3071bp + Sequence Listing; English.  
AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
amino acid sequences given in AAM82170 to AAM91921. (I) have cytoskeletal  
activity, and can be used in gene therapy and vaccine production. (I)

CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.

XX  
SQ Sequence 10579 BP; 2298 A; 2894 C; 3129 G; 2258 T; 0 other;

Query Match 26.7%; Score 414.4; DB 22; Length 10579;  
Best Local Similarity 81.6%; Pred. No. 1.1e-92;  
Matches 560; Conservative 0; Mismatches 1; Indels 125; Gaps 2;

QY 958 AGGTGTCATCAAGCTGAAGAACACGCGTATTCTTCATTCGCAATGAGGTCGACGGCC 1017  
DB 10005 AGGTGTCATCAAGCTGAAGAACACGCGTATTCTTCATTCGCAATGAGGTCGACGGCC 9946  
QY 1018 CATTCACATCGATGAGCGCGCGGTCTCTGTGGCTCCAAATGGCGCTCAGCAACACTC 1077  
DB 9945 CATTCACATCGATGAGCGCGCGGTCTCTGTGGCTCCAAATGGCGCTCAGCAACACTC 9886  
QY 1078 TGTGTGTG----- 1085  
DB 9885 TGTGTGTGAGGAGTCTGGGGAGAGGAGGCAAGATGAGACTGAGGCTGTGTGA 9826  
QY 1086 ----- 1085  
DB 9825 GCCAATACACCACTTGGTCTGTGGTCCACCTGTCTTAACCACTTGTCTCCACC 9766  
QY 1086 -AGATCGCAGCGCTGCGAATGCTTCTTATCAACGAGACCTATGCGCTCATCAGG 1144  
DB 9765 CAGATCGCAGCGCTGCGAATGCTTCTTATCAACGAGACCTATGCGCTCATCAGG 9706  
QY 1145 GCTGAGCTGCGCAAGATCACACACAGTAGAGATGTGTGCGAGACTGATGGGCCCTCTC 1204  
DB 9705 GCTGAGCTGCGCAAGATCACACACAGTAGAGATGTGTGCGAGACTGATGGGCCCTCTC 9646  
QY 1205 CGGCTGTTTCCCTGCGCACTCCAGCCCCCTTGAGCTGGGAACCTCAGGCTCTTGAAAAA 1264  
DB 9645 CGGCTGTTTCCCTGCGCACTCCAGCCCCCTTGAGCTGGGAACCTCAGGCTCTTGAAAAA 9586  
QY 1265 CCGTGGCAGTGGGAGGCTCAGCTGGGGCGCATTTGATTTGAGCTTTGAGGGAGATAGGG 1324  
DB 9585 CCGTGGCAGTGGGAGGCTCAGCTGGGGCGCATTTGATTTGAGGGAGATAGGG 9526  
QY 1325 CTGGCTTTGTTGAAGCGACAGAGGCTGAGAACTCAGGCTTCCCTAGATCCAGAGCCC 1384  
DB 9525 CTGGCTTTGTTGAAGCGACAGAGGCTGAGAACTCAGGCTTCCCTAGATCCAGAGCCC 9466  
QY 1385 TCCCATCTTCTCTCTTAAAAACAACCTAACCCCCCATTTACCCCCCATTTGCCACT 1444  
DB 9465 TCCCATCTTCTCTCTTAAAAACAAC-----CTACCCCCCATTTGCCACT 9418  
QY 1445 TCACCTCCGTGTCTCAGCTGATTAAGCCCTGACACTCTTTTATATGTTTCTTTTGA 1504  
DB 9417 TCACCTCCGTGTCTCAGCTGATTAAGCCCTGACACTCTTTTATATGTTTCTTTTGA 9358  
QY 1505 AATAAAGACACAGATTCCAAGTA 1530  
DB 9357 AATAAAGACACAGATTCCAAGTA 9332

RESULT 5  
ABA20015/C  
ID ABA20015 standard; DNA; 28995 BP.

XX ABA20015;  
AC 23-JAN-2002 (first entry)  
DT XX  
DE Human nervous system related polynucleotide SEQ ID NO 12346.  
XX  
XX Human; neurotropic; neuroprotective; cytosolic; dermatological; virocidic;  
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antisclerol; antianemic; antiarthritic; cancer;  
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200159063-A2.  
XX  
PD 16-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01334.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229347.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.



Db 19525 GCTGAGGCTGCCAAGATCAACACAGTAGAGAGTGTTGGCAGGACTCGTGGGCCCTCTC 19466  
Oy 1205 CGGCGTGTTCCTCCCTGACCACTCCAGCCCTTGTAGCTGGGAACCTCAGGCTCTGGAAAA 1264  
Db 19465 CGGCTGTTCCTCCCTGACCACTCCAGCCCTTGTAGCTGGGAACCTCAGGCTCTGGAAAA 19406  
Oy 1265 CCTGGGCACTGGGAGGCTCAGCTGGGGGCATTTGATTTGAGCGGAGAGATAGG 1324  
Db 19405 CTTGGGCACTGGGAGGCTCAGCTGGGGGCATTTGATTTGAGCGGAGAGATAGG 19346  
Oy 1325 CTGGCTTTTGTGAGACCGACGAGGCTGAGAACTCAGGCTTCCCTAGATCCAGCCCC 1384  
Db 19345 CTGGCTTTTGTGAGACCGACGAGGCTGAGAACTCAGGCTTCCCTAGATCCAGCCCC 19286  
Oy 1385 TCCCATCTTCTCTCTTAAACACCCCTACCCCATTCACCCCATTTGCACCT 1444  
Db 19285 TCCCATCTTCTCTCTTAAACACCC-----CTACCCCATTTGCACCT 19228  
Oy 1445 TCACCTCTGTGTCTCCAGCTGATTAGCCTCAGACTCTTTTATTGTTTCTTTGTA 1504  
Db 19237 TCACCTCTGTGTCTCCAGCTGATTAGCCTCAGACTCTTTTATTGTTTCTTTGTA 19178  
Oy 1505 AATAAAGACACAGGTTCCAAAGTA 1530  
Db 19177 AATAAAGACACAGGTTCCAAAGTA 19152

RESULT 6  
AAK70046/c  
ID AAK70046 standard; DNA; 28995 BP.  
XX  
AC AAK70046;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24858.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KM cyostatic; gene therapy; vaccine; metaatais; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0225219.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225216.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227189.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232387.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236357.  
PR 29-SEP-2000; 2000US-0236358.  
PR 29-SEP-2000; 2000US-0236359.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.

Query Match	26.7%;	Score 414.4;	DB 22;	Length 28995;
Best Local Similarity	81.6%;	Pred No. 1.6e-92;		
08-NOV-2000;	2000US-0246528.			
08-NOV-2000;	2000US-0246532.			
08-NOV-2000;	2000US-0246539.			
08-NOV-2000;	2000US-0246610.			
08-NOV-2000;	2000US-0246611.			
08-NOV-2000;	2000US-0246613.			
17-NOV-2000;	2000US-0249207.			
17-NOV-2000;	2000US-0249208.			
17-NOV-2000;	2000US-0249209.			
17-NOV-2000;	2000US-0249210.			
17-NOV-2000;	2000US-0249211.			
17-NOV-2000;	2000US-0249212.			
17-NOV-2000;	2000US-0249213.			
17-NOV-2000;	2000US-0249214.			
17-NOV-2000;	2000US-0249215.			
17-NOV-2000;	2000US-0249216.			
17-NOV-2000;	2000US-0249217.			
17-NOV-2000;	2000US-0249218.			
17-NOV-2000;	2000US-0249244.			
17-NOV-2000;	2000US-0249245.			
17-NOV-2000;	2000US-0249264.			
17-NOV-2000;	2000US-0249265.			
17-NOV-2000;	2000US-0249297.			
17-NOV-2000;	2000US-0249299.			
17-NOV-2000;	2000US-0249300.			
01-DEC-2000;	2000US-0250160.			
01-DEC-2000;	2000US-0250391.			
05-DEC-2000;	2000US-0251030.			
05-DEC-2000;	2000US-0251988.			
05-DEC-2000;	2000US-0256719.			
06-DEC-2000;	2000US-0251479.			
08-DEC-2000;	2000US-0251856.			
08-DEC-2000;	2000US-0251868.			
08-DEC-2000;	2000US-0251869.			
08-DEC-2000;	2000US-0251989.			
08-DEC-2000;	2000US-0251990.			
11-DEC-2000;	2000US-0254097.			
05-JAN-2001;	2001US-0259678.			
(HUMA-) HUMAN GENOME SCI INC.				
PA				
PI	Rosen CA, Barash SC, Ruben SM;			
XX				
DR	WPI; 2001-483426/52.			
XX				
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,			
XX	useful for preventing, diagnosing and/or treating cancers and			
PT	metastasis -			
XX				
PS	Disclosure; SEQ ID NO 24858; 3071pp + Sequence Listing; English.			
XX				
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)			
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytotoxic			
CC	activity, and can be used in gene therapy and vaccine production. (I)			
CC	proteins and polynucleotides may be used in the prevention, diagnosis and			
CC	treatment of diseases associated with inappropriate (I) expression. For			
CC	example, they may be used to treat disorders associated with decreased			
CC	expression by rectifying mutations or deletions in a patient's genome			
CC	that affect the activity of (I) by expressing inactive proteins or to			
CC	supplement the patient's own production of (I). Additionally, (I)			
CC	polynucleotides may be used to produce the secreted (I), by inserting			
CC	the nucleic acids into a host cell and culturing the cell to express the			
CC	protein. (I) proteins and polynucleotides may be used to prevent,			
CC	diagnose and treat immune/hematopoietic-related diseases, especially			
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703			
CC	to AAK57694 represent human immune/hematopoietic antigen genomic			
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169			
CC	represent sequences used in the exemplification of the present invention.			
XX				
XX				
XX	Sequence 28995 BP; 6151 A; 8327 C; 8051 G; 6466 T; 0 other;			

[illegible]



PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225256.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231245.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231966.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234999.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249267.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251858.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-483426/52.  
 XX Nucleic acid encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX

PS Disclosure; SEQ ID NO 34779; 3071bp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytosolic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK67694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM62169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 28995 BP; 6466 A; 8051 C; 8327 G; 6151 T; 0 other;  
  
Query Match 26.7%; Score 414.4; DB 22; Length 28995;  
Best Local Similarity 81.6%; Pred. No. 1.6e-92;  
Matches 560; Conservative 0; Mismatches 1; Indels 125; Gaps 2;  
  
QY 958 AGGTGTCATCAAGCTGAAAGAACACGCGTATTTCTTCATTGCGCAAGAGGTCGACGGCC 1017  
DB 9171 AGGTGTCATCAAGCTGAAAGAACACGCGTATTTCTTCATTGCGCAAGAGGTCGACGGCC 9230  
  
QY 1018 CATTCATCATGATGAGGCGCGGCTCTGTGGCTCCAAATGGCGGCTCAGCAACAATC 1077  
DB 9231 CATTCATCATGATGAGGCGCGGCTCTGTGGCTCCAAATGGCGGCTCAGCAACAATC 9290  
  
QY 1078 TGTGTGTG----- 1085  
DB 9291 TGTGTGTGAGTGTAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9350  
QY 1086 ----- 1085  
DB 9351 GCCAGTACACCACTTGTCTGTGTCACCTCTGTCTTAAGCCACTTTGTCTCCACC 9410  
QY 1086 -AGATGCGAGCGCTGATTCGTTCTTATCAACAGGAGCTCATGAGCTCATCAGG 1144  
DB 9411 CAGATGCGAGCGCTGATTCGTTCTTATCAACAGGAGCTCATGAGCTCATCAGG 9470  
  
QY 1145 GCTGAGGCTGCAAGATCACCAACAGTGAAGATGTGGCAGAGCTCGTGGCCCTCTC 1204  
DB 9471 GCTGAGGCTGCAAGATCACCAACAGTGAAGATGTGGCAGAGCTCGTGGCCCTCTC 9530  
  
QY 1205 CGGCTGTTTCCCTGCTCCAGCTCCAGCCCTTGAAGCTGGGAATCAGGCTCCTGGAAAAA 1264  
DB 9531 CGGCTGTTTCCCTGCTCCAGCTCCAGCCCTTGAAGCTGGGAATCAGGCTCCTGGAAAAA 9590  
  
QY 1265 CTTGGGCACTGGGAGGCTCAGCTGGGCGCATGATTTGAGCTTTGAGGAGAGATAGGG 1324  
DB 9591 CTTGGGCACTGGGAGGCTCAGCTGGGCGCATGATTTGAGCTTTGAGGAGAGATAGGG 9650  
  
QY 1325 CTGGCTTTGTGAAGCAGACAGAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCC 1384  
DB 9651 CTGGCTTTGTGAAGCAGACAGAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCC 9710  
  
QY 1385 TCCCATCTTCTCTCTCTTAAACAAACCTACCCCATTTCAACCCCATTTGCACT 1444  
DB 9711 TCCCATCTTCTCTCTCTTAAACAAACCTACCCCATTTCAACCCCATTTGCACT 9758  
  
QY 1445 TCAGCTCTGCTGCTCAGGCTGATTAAGCTCAGACTCTTTATTTATTTTCTTTTGA 1504  
DB 9759 TCAGCTCTGCTGCTCAGGCTGATTAAGCTCAGACTCTTTATTTATTTTCTTTTGA 9818  
  
QY 1505 AATAAAGACACAGGCTCCAAAGTA 1530  
DB 9819 AATAAAGACACAGGCTCCAAAGTA 9844

RESULT 8  
ID AAK65213  
AAK65213 standard; DNA; 28995 BP.  
AC AAK65213;  
XX  
XX 07-NOV-2001 (first entry)  
DT  
XX  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40025.  
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
XX Homo sapiens.  
OS  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0188874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 09-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226581.  
PR 22-AUG-2000; 2000US-0226586.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0231080.

PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234224.  
PR 25-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251859.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruden SM;  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and  
XX metacasis -  
XX  
XX Disclosure; SEQ ID NO 40025; 3071bp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
XX activity, and can be used in gene therapy and vaccine production. (I)  
XX proteins and polynucleotides may be used in the prevention, diagnosis and  
XX treatment of diseases associated with inappropriate (I) expression. For  
XX example, they may be used to treat disorders associated with decreased  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of (I) by expressing inactive proteins or to  
XX supplement the patient's own production of (I). Additionally, (I)  
XX polynucleotides may be used to produce the secreted (I), by inserting the  
XX nucleic acids into a host cell and culturing the cell to express the  
XX protein. (I) proteins and polynucleotides may be used to prevent,  
XX diagnose and treat immune/hematopoietic-related diseases, especially  
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
XX to AAK87694 represent human immune/hematopoietic antigen genomic  
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
XX represent sequences used in the exemplification of the present invention.  
SQ Sequence 28995 BP; 6466 A; 8051 C; 8327 G; 6151 T; 0 other;  
Query Match 26.7%; Score 414.4; DB 22; Length 28995;  
Best Local Similarity 81.6%; Pred. No. 1.6e-92;  
Matches 560; Conservative 0; Mismatches 1; Indels 125; Gaps 2;  
Qy 958 AGGTGTCATCAAGCTGAAGAACACGCGTATTTCTTCAATGAGGTCGACCGGCC 1017  
Db 9171 AGGTGTCATCAAGCTGAAGAACACGCGTATTTCTTCAATGAGGTCGACCGGCC 9230  
Qy 1018 CATCTACATCATGAGACGCGCGGTGCTCTGTGCTCCAAATGAGCGCTCAGACCAACTC 1077  
Db 9231 CATCTACATCATGAGACGCGCGGTGCTCTGTGCTCCAAATGAGCGCTCAGACCAACTC 9290  
Qy 1078 TGTGCTGG----- 1085  
Db 9291 TGTGCTGAGGTGAGCTGGGAGAGGAGGACGAGACCTAGGCGTGTGA 9350  
Qy 1086 ----- 1085  
Db 9351 GCCAGTACACCAACTGTCCTGTGTCACCTGTCTTAAGCCACTTGTCTCCACC 9410  
Qy 1086 -AGATGCGCAGCTGCGATGCTGCTTCTTATCAACGAGACTTATGCTCCATCAGG 1144  
Db 9411 CAGATGCGCAGCTGCGATGCTGCTTCTTATCAACGAGACTTATGCTCCATCAGG 9470  
Qy 1145 GCTGAGGCTGCCAAGATACACCAAGTGAAGAAATGTTGAGAGACTGTCGCGCTCTTC 1204  
Db 9471 GCTGAGGCTGCCAAGATACACCAAGTGAAGAAATGTTGAGAGACTGTCGCGCTCTTC 9530

Qy	1205	CGCTGTTTCCCTGCACTCCAGCCCCCTTAGCTGGAACTCAGAGCTCTGSAAAA	1264
Db	9551	CGGCTGTTTCCCTGCACTCCAGCCCTTAGCTGGAACTCAGAGCTCTGSAAAA	9590
Qy	1265	CCTGGGCAGTGGAGGCTCAGCTGCGGGCCATTGATTGAGCTTTGAGGGAGATAGG	1324
Db	9551	CTTGGGCAGTGGAGGCTCAGCTGCGGGCCATTGATTGAGCTTTGAGGGAGATAGG	9550
Qy	1325	CTGGCTTTGTGAAGCCAGCAGAGGCTGAGAACCTCAGGCTTCCTAGATCCAGGCC	1384
Db	9651	CTGGCTTTGTGAAGCCAGCAGAGGCTGAGAACCTCAGGCTTCCTAGATCCAGGCC	9710
Qy	1385	TCCCATTCCTCTCTCTTAAAAACAACCTCAGCCCACTTACCCCCCATTCGCACT	1444
Db	9711	TCCCATTCCTCTCTCTTAAAAACAAC-----CTACCCCACTTCGCACT	9758
Qy	1445	TCACTCTGTGTCTCCAGCTGATTAAGCTTCTTTATTTGTTTCTTTTGT	1504
Db	9759	TCACTCTGTGTCTCCAGCTGATTAAGCTTCTTTATTTGTTTCTTTTGT	9818
Qy	1505	AATAAAGACACAGGTTCCAAAGTA	1530
Db	9819	AATAAAGACACAGGTTCCAAAGTA	9844

## RESULT 9

ID ABLO2115 standard; cDNA; 1905 BP.  
 XX  
 AC ABLO2115;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 827.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN MO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers BW;  
 XX  
 DR WPI: 2001-656660/75.  
 DR P-PSDB; ABB58012.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 XX interactions -  
 XX  
 PS Claim 1, SEQ ID NO 827; 21pp + Sequence Listing; English.  
 CC  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
 CC sequences (ABLO1840-ABU16175) and the encoded proteins  
 CC (ABB57737-ABBY2072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO

CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

Sequence 1905 BP; 490 A; 568 C; 466 G; 381 T; 0 other.

Query Match 22.6%; Score 351.4; DB 23; Length 1905;

Matches 609; Conservative 0; Mismatches 386; Indels 6; Gaps 2;

Qy	166	GGTGAACCAAGATCTGGGCGCGTGGAAAGCTGGAAATGAACCTCCGTGCATTAATATGCGT	225
Db	786	GCCACCAAGATTTGGGCGCGTGGAAAGCCCATTTGACGCTTGGCCCTCATATATTTGGCAT	845
Qy	226	GTTGCAACCAACCACTGACCTCCGTCCACCTGGACCTGGGAATTCAGCTGCCGCTTAC	285
Db	846	CCAGCAACCAACCATCTGGCGGATTAATCCATCCGCGCGTTAAGTTTTCTCGAAGTTTAC	905
Qy	286	CTTTTCGGAGGTCAGAGAGCTTTGGTATGCGCCCTGCTCTTACATCTCTGTCACTTCAAGTT	345
Db	906	TCTGAGGAATCTACAGACGCGCTGTATGTCTGTGTATGAGCTCGGCGGTCCAGAT	965
Qy	346	GGCCTGTCAAGCCCATGAGGAGCTGACCCCAAGAGGTATTTGAGGCATTCAGCAAGC	405
Db	966	TGCAGTGTCCGCCTAATGAAGACTTGACCCCACTCGTGAGTCCGTTCAAGCCAAAGC	1025
Qy	406	CCTGTTTGAAGAGCTGAGAGCAGCTGTGTGAACAAAGTGGATCGACCAAGCCAC	465
Db	1026	TTTGTACAGCGCTGAGAGGAGGATCTGCTTGGCACATTAAGACTCGGAACCAACAA	1085
Qy	466	CTTGGAGACTTCCAGGACCTGCTGCAACAGACCTGTATGCTTCTTACCTGGCCGTAC	525
Db	1086	ACTTGAGCAGTTTCCAGAGCTTACCTGTGTAAGATGTCTCCGTTTCTACTGGCCCGAC	1145
Qy	526	CGCGAAGGCCCTGACGGCCCACTGGAGCTCATGAAAGCATTAACCTGTGAGAGCA	585
Db	1146	TGCCAATCTTGTCAAAACCATTTGGCTTTCTTCAAGCACTACACCTGTTGCCGATCA	1205
Qy	586	GACAGTGCAGCCGCT--GCCCAAGGGAGCAAGTGTGAACCTTCTGTATGCAGAGA	642
Db	1206	GTCAAGTAAAGCTTATATATGAGCAAGGATCAGACAGCCACTCAGCTTCTCGAATGCAGAGA	1265
Qy	643	CCTGATTTGATGACATGATAGCTCAAGGACATGCGAGATGAGGTCTTGGAACTGAGCTAT	702
Db	1266	TCAATCTTTCAGACACGACTTGACAGCCACTGTGCGAAGCTCTCGAGATGAGACGAGC	1325
Qy	703	GSTGCTGACCGCGCGCCAGAGAGAGATTTGGGACGTGGAACAGAACTGCATTAATG	762
Db	1326	TCTGGCAGATGTGTGCAACCAAGTAATATGCTGCTGTGGAGAAAGAGCTCTCGCGTG	1385
Qy	763	GCAAGTGTATGTGACAGCATCAAGGCAATGAGCTTCCG--GACTTGCACACCAAGAC	819
Db	1386	GACCGTCTGTGTGATTTCCGTGCTTATGTCCGACCGCTGCTCCGAGTTGCACACCAAGAC	1445
Qy	820	ACTGGCACTGCTGCGGGGCCGCAATGTTGGGTATCTGATATGCGCTCGCGTGAATCAACCT	879
Db	1446	ACTGGCTCTGTTGTGTGGCCGCAATGTGGGTATCTTAATGCTTCCAAAGAAATTAACCTT	1505
Qy	880	GGGCAAGACCAACCAAGATTAACCAAGATTGATGGAACCTGTCTTGAAGGGTCCGCGCTG	939
Db	1506	TGTTCCGATGCCAAGACCTGTGTGTGACGTGTGATCTGGGGCTCGAGGACACCGGCTGC	1565
Qy	940	GAAAGTATCCCGAAMCAAGGTGTCAATCAAGCTGAAAGAACAGGTGATTTCTTCATTGC	999
Db	1566	GAAAGTCTCTGCGCGCAAGAACCTTTAAGCTGCGACCAATGGGGATTTCTTCATGCG	1625
Qy	1000	CAATGAGGTTCGACGGCCCATCTACATTCATGATGACCGCGCGTCTGTGGTCCAAATG	1059
Db	1626	TAAACAGGAAAGAGGCACTTTTATCGACGCACTCTTTGTATTCGCCAAACAAAGC	1685
Qy	1060	GCGCTCAGCAACAACCTGTGTGATGAGATTCGCAAGCTCGCAATTCGTTCTTATCA	1119
Db	1686	TGCACTCGGTACAACTGACACATGGAATATTCGCTCGCCCTTCAACCTTCTGTGTAA	1745
Qy	1120	CGAGACTCATTTGCCCTCATCAAGGCTGAAGCTGCCAAGA	1160



```
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226881.
PR 22-AUG-2000; 2000US-0226886.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0255678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
```

PI Rosen CA, Barash SC, Ruben SM;  
XX WPI, 2001-483426/52.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
PS Disclosure; SEQ ID NO 34778; 3071bp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 8204 BP; 1865 A; 2157 C; 2229 G; 1953 T; 0 other;  
Query Match 10.3%; Score 159.4; DB 22; Length 8204;  
Best Local Similarity 99.4%; Pred. No. 3e-29; Mismatches 0; Gaps 0;  
Matches 160; Conservative 0; Indels 1;  
QY 72 AGGATCCAAAGCCCCGACACTCTGTGCCACCCAGCCCGCCCTGGACTCA 131  
Db 3245 AGGATCCAAAGCCCCGACACTCTGTGCCACCCAGCCCGCCCTGGACTCA 3304  
QY 132 CCAAGCGGTGAAGAAGTAAACAGCCACTTCAGGTGACCAAGATCTGGGCGCTGGA 191  
Db 3305 CCAAGCGGTGAAGAAGTAAACAGCCACTTCAGGTGACCAAGATCTGGGCGCTGGA 3364  
QY 192 AGCCTGCAATGACTCTCTCATTAATGCTGTGTGCG 232  
Db 3365 AGCCTGCAATGACTCTCTCATTAATGCTGTGTGCG 3405  
RESULT 12  
AAK85212  
ID AAK85212 standard; DNA; 8204 BP.  
XX  
AC AAK85212;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO:40024.  
XX  
KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
FN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218230.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236357.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249399.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX Disclosure; SEQ ID NO 40024; 3071pp + Sequence Listing; English.  
XX  
XX AAK64951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic

CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK67694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 8204 BP, 1865 A; 2157 C; 2229 G; 1953 T; 0 other;  
  
Query Match 10.3%; Score 159.4; DB 22; Length 8204;  
Best Local Similarity 99.4%; Pred. No. 3e-29; Mismatches 0; Gaps 0;  
Matches 160; Conservative 0; Indels 1;  
  
QY 72 AGGATCCAAAGCCCGCAGCACTCTGTGCCACCCAGCCCGCAGCCCTGACTCA 131  
DB 3245 AGGTATCCAAAGCCCGCAGCACTCTGTGCCACCCAGCCCGCAGCCCTGACTCA 3304  
QY 132 CCNAGCGTGTGAAGAAGTAAACAGCCACTTCAGGTGACCAAGATCTGGCGCGCTGGA 191  
DB 3305 CCNAGCGTGTGAAGAAGTAAACAGCCACTTCAGGTGACCAAGATCTGGCGCGCTGGA 3364  
QY 192 AGCCTGCAATGACCTCTGCTCATTAATGCTGTGTTGCA 232  
DB 3365 AGCCTGCAATGACCTCTGCTCATTAATGCTGTGTTGCA 3405  
  
RESULT 13  
AAT23953  
ID AAT23953 standard; cDNA to mRNA; 144 BP.  
XX  
XX AAT73953;  
AC  
XX  
XX 27-AUG-1996 (first entry)  
DT  
XX  
XX Human gene signature HUMG05899.  
DE  
XX  
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KW human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W09514772-A1.  
PN  
XX  
XX 01-JUN-1995.  
PD  
XX  
XX I1-NOV-1994; 94WO-JP01916.  
PF  
XX  
XX 12-NOV-1993; 93JP-0355504.  
PR  
XX  
XX (MATS/) MATSUBARA K.  
PA (OKUB/) OKUBO K.  
XX  
XX  
PI Matsubara K, Okubo K;  
XX  
XX WPI; 1995-206931/27.  
DR  
XX  
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissues  
XX  
XX  
XX Claim 1; Page 1495; 2245pp; Japanese.



CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC given in AAT19001-r26837 and which is able to hybridise to part of  
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared from  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.

XX Sequence 144 BP; 34 A; 50 C; 15 G; 44 T; 1 other;  
SQ

Query Match 7.8%; Score 121; DB 16; Length 144;  
Best Local Similarity 91.7%; Pred. No. 2e-20;  
Matches 143; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

Oy 1372 GATCCAGAGCCCCCTCCCATCTTCTCTCTTAAACACACCCCTACCCCATTTACCC 1431  
Db 1 GATCCAGAGCCCCCTCCCATCTTCTCTCTTAAACACAC-----CTACCC 48

Oy 1432 CCGATTGCCACCTTCACCTCGTCTCCAGCTGATTAGCCCTACACTTCTTATTG 1491  
Db 49 CCGATTGCCACCTTCACCTCGTCTCCAGCTGATTAGCCCTACACTTCTTATTG 108

Oy 1492 TTTTCTTTTGAATAAACACACAGGTTCCAAA 1527  
Db 109 TTTTCTTTTGAATAAACACACAGGTTCCAAA 144

RESULT 14  
AAK87783  
ID AAK87783 standard; cDNA; 457 BP.  
XX AAK87783;  
AC  
XX  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human digestive system antigen coding sequence SEQ ID NO: 99.  
XX  
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;  
XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
XX digestive system disorder; Meckel's diverticulum; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX WO200155314-A2.  
XX  
XX  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01324.  
XX  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236337.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.

PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM,	
XX	MPI: 2001-502630/55.	
XX	P-PSDB; AAM92010.	
XX	Polynucleotides encoding digestive system antigens, useful for	
XX	diagnosing, treating, preventing and/or prognosing disorders of the	
XX	digestive system, particularly cancer and cancer metastases -	
XX	Claim 1; SEQ ID NO 99; 986pp; English.	
XX	The present invention provides the protein and coding sequences of a	
XX	number of human digestive system antigens. These can be used in the	
XX	diagnosis, treatment and prevention of digestive system disorders,	
XX	including cancer, Meckel's diverticulum, bacterial or parasitic	
XX	infections, appendicitis, Hirschsprung's disease, chronic colitis or	
XX	ulcerative colitis. The present sequence is a cDNA encoding a digestive	
XX	system antigen of the invention.	
XX	Sequence 457 BP; 94 A; 129 C; 123 G; 109 T; 2 other;	
XX	Query Match 4.6%; Score 71.6; DB 22; Length 457;	
XX	Best Local Similarity 94.4%; Pred. NO. 6:5e-08;	

[illegible]

PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0226924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234224.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235835.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246605.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249219.  
 PR 17-NOV-2000; 2000US-0249220.  
 PR 17-NOV-2000; 2000US-0249221.  
 PR 17-NOV-2000; 2000US-0249222.  
 PR 17-NOV-2000; 2000US-0249223.  
 PR 17-NOV-2000; 2000US-0249224.  
 PR 17-NOV-2000; 2000US-0249225.  
 PR 17-NOV-2000; 2000US-0249226.  
 PR 17-NOV-2000; 2000US-0249227.  
 PR 17-NOV-2000; 2000US-0249228.  
 PR 17-NOV-2000; 2000US-0249229.  
 PR 17-NOV-2000; 2000US-0249230.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 P1 Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-457728/49.  
 XX  
 DR P-PSDB; AAU19989.  
 DR  
 XX  
 PT Isolated nucleic acid molecule encoding a human liver related protein  
 PT is used in preventing, treating or ameliorating disorders of the liver  
 PT particularly cancer of the liver -  
 XX  
 PS Claim 1; SEQ ID NO 30; 526bp; English.  
 XX  
 CC Sequences AA31681-AA31826 represent cDNA molecules, which encode the  
 CC liver associated polypeptides of the invention. Liver associated  
 CC polypeptides and their associated polynucleotides are useful in the  
 CC diagnosis, treatment and prevention of various types of disorders in e.g.  
 CC human, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A  
 CC pathological condition can be determined by detecting the presence or  
 CC absence of a mutation in a liver associated polynucleotide. The treatable  
 CC disorders include autoimmune diseases such as rheumatoid arthritis,  
 CC hyperproliferative disorders such as neoplasms of the breast or liver,  
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular  
 CC disorders such as cerebral ischemia, nervous system disorders such as  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,  
 CC ocular disorders such as corneal infection, endocrine disorders such as  
 CC premature labour and infertility, gastrointestinal disorders such as  
 CC Crohn's disease, renal disorders such as glomerulonephritis and  
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can  
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,  
 CC to maintain organs before transplantation, to regenerate tissues and in  
 CC chemotaxis.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 Query Match 4.6%; Score 71.6; DB 22; Length 457;  
 Best Local Similarity 94.4%; Pred. No. 6.5e-08;  
 Matches 85; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 Oy 958 AGGTGTCATCAAGCTGAAGACAAGGTGATTTCTTCATTCGCAATGAGGTGACGCGC 1017

Db 276 AGGTGTCATCAAGCTGAGAGACACGGTGATTTCTTCATTGCCAATGAGGGTCGACGG-C 334

Oy 1018 CATCTACATCGATGAGCGGCCGGTCTCTG 1047

Db 335 CATCTACATCGATGAGCGGCCGGTCTCTCTG 364

Search completed: March 27, 2003, 05:40:40  
Job time : 632 secs

• • •

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 05:29:20 ; Search time 81 Seconds  
(without alignments)  
5879.864 Million cell updates/sec

Title: US-09-744-125A-1  
Perfect score: 1553  
Sequence: 1 gaattcgccagcagtgctggg.....aaaaaaaaaacctcag 1553

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.4	3.2	1926	US-09-249-585A-2	Sequence 2, Appli
2	49.4	3.2	2580	US-09-050-863-2	Sequence 2, Appli
3	49.4	3.2	2580	US-09-359-081-2	Sequence 2, Appli
4	49.4	3.2	5452	US-09-130-114-1	Sequence 1, Appli
5	49.4	3.2	9600	US-08-910-647-1	Sequence 1, Appli
6	49.4	3.2	9600	US-09-620-925-1	Sequence 1, Appli
7	49.4	3.2	10596	US-07-884-811-15	Sequence 15, Appli
8	49.4	3.2	10596	US-07-885-971-15	Sequence 15, Appli
9	49.4	3.2	10596	US-08-087-783A-15	Sequence 15, Appli
10	49.4	3.2	10596	US-08-194-088B-15	Sequence 15, Appli
11	49.4	3.2	10596	US-08-194-087-15	Sequence 15, Appli
12	49.4	3.2	10596	PCT-US93-04648-15	Sequence 15, Appli
13	49.4	3.1	3489	US-08-728-323A-1	Sequence 1, Appli
14	48.6	3.1	3489	US-09-298-568-1	Sequence 1, Appli
15	48.6	3.1	32207	US-08-770-379-20	Sequence 20, Appli
16	48.6	3.1	32207	US-08-757-669A-20	Sequence 20, Appli
17	48.6	3.1	32207	US-09-230-371A-20	Sequence 20, Appli
18	47.6	3.1	16442	US-08-781-891-208	Sequence 208, App
19	46	3.0	7218	US-08-232-463-14	Sequence 14, Appli
20	45.6	2.9	913	US-08-217-327-3	Sequence 3, Appli
21	45.6	2.9	913	US-07-885-970A-3	Sequence 3, Appli
22	45.6	2.9	913	US-08-298-687A-3	Sequence 3, Appli
23	45.6	2.9	913	US-08-530-797-2	Sequence 2, Appli
24	45.6	2.9	913	US-08-298-829-3	Sequence 3, Appli
25	45.6	2.9	913	US-08-787-335-2	Sequence 2, Appli
26	45.6	2.9	1984	US-07-885-970A-25	Sequence 25, Appli
27	45.6	2.9	1985	US-08-298-687A-25	Sequence 25, Appli

C	28	45.6	2.9	1985	1	US-08-298-829-25	Sequence 25, Appli
C	29	45.6	2.9	7218	1	US-08-232-463-14	Sequence 14, Appli
C	30	42.6	2.7	2338	1	US-08-425-069-1	Sequence 1, Appli
C	31	42.6	2.7	2338	2	US-08-317-844B-1	Sequence 1, Appli
C	32	42.6	2.7	2793	1	US-08-209-747-1	Sequence 1, Appli
C	33	42	2.7	2793	1	US-08-458-298-1	Sequence 1, Appli
C	34	41.4	2.7	397	3	US-09-253-691-3	Sequence 3, Appli
C	35	41	2.6	432	1	US-08-642-255-48	Sequence 48, Appli
C	36	41	2.6	756	1	US-08-642-255-50	Sequence 50, Appli
C	37	39.8	2.6	51259	3	US-08-781-891-209	Sequence 209, App
C	38	39.6	2.5	1508	4	US-09-039-046-1	Sequence 1, Appli
C	39	39.6	2.5	5267	3	US-08-976-255-2	Sequence 2, Appli
C	40	38.8	2.5	2830	2	US-09-010-928B-1	Sequence 1, Appli
C	41	38.8	2.5	9551	1	US-08-056-200-93	Sequence 93, Appli
C	42	38.8	2.5	9551	2	US-08-800-644-93	Sequence 93, Appli
C	43	38.4	2.5	9757	1	US-08-093-453B-1	Sequence 1, Appli
C	44	38.4	2.5	9759	1	US-08-459-041A-1	Sequence 1, Appli
C	45	38.4	2.5	9759	3	US-08-999-733-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-249-585A-2  
; Sequence 2, Application US/09249585A  
; Patent No. 6417002  
; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert  
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES  
; FILE REFERENCE: 0867/0D905  
; CURRENT APPLICATION NUMBER: US/09/249,585A  
; CURRENT FILING DATE: 1999-02-11  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 1926  
; TYPE: DNA  
; ORGANISM: Epstein Barr Virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1926)  
; OTHER INFORMATION: coding strand of EBNA-1 DNA  
US-09-249-585A-2

Query Match 3.2%; Score 49.4; DB 4; Length 1926;  
Best Local Similarity 49.8%; Pred. No. 0.0015;  
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY	527	GCGAAGGCGCTGCGAGGCCACCTGACGCTCAATGAAGCAGTATTACTGCGAGGACCG	586
DB	359	GCGCAGAGGGGCGAGCAGCAGGAGGCGCAGAGCGAGGCGCAGAGGCGCAGAGGCGCAG	418
QY	587	ACAGTGCAGCCGCTGCCCAAGAGGAGCCCAAGTCTGAATCTTCTGATGACAGGACCTG	646
DB	419	GAGGGGCGAGAGCAGCAGAGGAGGCGCAGAGCAGAGGAGGCGCAGAGGCGCAGAGCAG	478
QY	647	ATTGATGACAGTACCTAAGCAGCAGTGCAGATGAGTCTTGGAACATGAGCTGATGTG	706
DB	479	GAGGAGGGGCGAGAGGGGCGAGGCGCAGAGCAGCAGGAGGAGGCGCAGAGCAGAGGAG	538
QY	707	GCTGCGCGGCGCGAAGCAGAGATTGCGCAGCTGGAACGAACCTGATTAAGTGGCG	766
DB	539	GCGCAGAGGGGCGAGCAGAGGAGGCGCAGAGGCGCAGAGGCGCAGAGGCGCAGAGGAG	598
QY	767	GTGCTAGTGGGA 777	
DB	599	GAGGGGCGAGGA 609	

RESULT 2  
US-09-050-863-2  
; Sequence 2, Application US/09050863

```
Patent No. 6114111
GENERAL INFORMATION:
APPLICANT: Lao, Ying
APPLICANT: Hiang, Betty
APPLICANT: Payan, Don
TITLE OF INVENTION: Mammalian Protein Interaction Cloning
TITLE OF INVENTION: System
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,863
FILING DATE: 30-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 949-8711
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-09-050-863-2

Query Match          3.2%; Score 49.4; DB 3; Length 2580;
Best Local Similarity 49.8%; Pred. No. 0.0017;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGGCGCCCTGCGAGCCCACTGCGACATGACAGTATTACTGCTGAGAGCCAG 586
DB 742 GGGCAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
QY 587 ACAGTGCAGCCCGCTGCCAAGAGGAGCAAGTGTGAACCTTCTGTATGCAAGACCTG 646
DB 802 GAGGGGCGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
QY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGAACATGAGCTGATG 706
DB 862 GAGGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
QY 707 GCTGACCGGCGCCAGAAAGGAGAGATTCGCGACCTGGAACAGGAATCTGATAAGTGC 766
DB 922 GGGCAGAGAGGGGCGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
QY 767 GTGCTAGTGA 777
DB 982 GAGGGGCGAGGA 992
```

```
RESULT 3
US-09-359-081-2
Sequence 2, Application US/09359081
```

```
GENERAL INFORMATION:
APPLICANT: Lao, Ying
APPLICANT: Hiang, Betty
APPLICANT: Payan, Don
```

```
TITLE OF INVENTION: Mammalian Protein Interaction Cloning
System
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/359,081
FILING DATE: 22-Jul-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/050,863
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 949-8711
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match          3.2%; Score 49.4; DB 4; Length 2580;
Best Local Similarity 49.8%; Pred. No. 0.0017;
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGGCGCCCTGCGAGCCCACTGCGACATGACAGTATTACTGCTGAGAGCCAG 586
DB 742 GGGCAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
QY 587 ACAGTGCAGCCCGCTGCCAAGAGGAGCAAGTGTGAACCTTCTGTATGCAAGACCTG 646
DB 802 GAGGGGCGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
QY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGAACATGAGCTGATG 706
DB 862 GAGGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
QY 707 GCTGACCGGCGCCAGAAAGGAGAGATTCGCGACCTGGAACAGGAATCTGATAAGTGC 766
DB 922 GGGCAGAGAGGGGCGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
QY 767 GTGCTAGTGA 777
DB 982 GAGGGGCGAGGA 992
```

```
RESULT 4
US-09-130-114-1/c
Sequence 1, Application US/09130114
```

```
GENERAL INFORMATION:
APPLICANT: Horlick, Robert A.
APPLICANT: Dama, Bassam B.
APPLICANT: Robbins, Alan K.
TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
```

TITLE OF INVENTION: From Multiple Transfected Epismes  
FILE REFERENCE: 0867/1D903US1  
CURRENT APPLICATION NUMBER: US/09/130.114  
CURRENT FILING DATE: 1998-08-06  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 5452  
TYPE: DNA  
ORGANISM: VEBNA  
US-09-130-114-1

Query Match 3.2%; Score 49.4; DB 2; Length 5452;  
Best Local Similarity 49.8%; Pred. No. 0.0025;  
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

OY 527 GCGAAGCCCTGCAGGCCCTGAGCTGATGAAGCATATTACTGTGAGAGACCG 586  
DB 2063 GGGCAGAGGGGCGAGAGCAGAGAGGGGCGAGAGAGAGAGAGAGAGAG 2004  
OY 587 ACAGTGCAGCCGCTGCCCAAGAGGAGCCAAAGTGTGAATCTTGATGACAGAGCCTG 646  
DB 2003 GAGGGGCGAGAGCAGAGAGAGGGGCGAGAGAGGGGCGAGAGAGAGAG 1944  
OY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTTGGAACATGAGCTGATG 706  
DB 1943 GAGGAGGGGCGAGAGAGGGGCGAGAGAGGGGCGAGAGAGAGAGAGAG 1884  
OY 707 GCTGACCGGCGCCAGAGAGAGATTGGCAGCTGGAAACAGGAAGTGCATTAAGTGCGAG 766  
DB 1883 GGGCAGAGGGGCGAGAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGAGAG 1824  
OY 767 GTGCTAGTGA 777  
DB 1823 GAGGGGCGAGAG 1813

## RESULT 5

US-08-910-647-1  
Sequence 1, Application US/08910647  
Patent No. 6251433  
GENERAL INFORMATION:  
APPLICANT: Zuckermann et al.  
TITLE OF INVENTION: Compositions and Methods for  
TITLE OF INVENTION: Polynucleotide Delivery  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910.647  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9600 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-910-647-1

Query Match 3.2%; Score 49.4; DB 4; Length 9600;  
Best Local Similarity 49.8%; Pred. No. 0.0033;  
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

OY 527 GCGAAGCCCTGCAGGCCCTGAGCTGATGAAGCATATTACTGTGAGAGACCG 586  
DB 788 GGGCAGAGGGGCGAGAGCAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAG 847  
OY 587 ACAGTGCAGCCGCTGCCCAAGAGGAGCCAAAGTGTGAATCTTGATGACAGAGCCTG 646  
DB 848 GAGGGGCGAGAGCAGAGAGAGGGGCGAGAGAGGGGCGAGAGAGAGAGAGAG 907  
OY 647 ATTGATGACAGTAAGCTCAAGACATGCGAGATGAGTCTTGGAACATGAGCTGATG 706  
DB 908 GAGGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAG 967  
OY 707 GCTGACCGGCGCCAGAGCAGAGATTGGCAGCTGGAAACAGGAAGTGCATTAAGTGCGAG 766  
DB 968 GGGCAGAGGGGCGAGAGCAGAGAGGGGCGAGAGAGGGGCGAGAGAGAGAGAG 1027  
OY 767 GTGCTAGTGA 777  
DB 1028 GAGGGGCGAGAG 1018

## RESULT 6

US-09-620-925-1  
Sequence 1, Application US/09620925  
Patent No. 6468986  
GENERAL INFORMATION:  
APPLICANT: Zuckermann et al.  
TITLE OF INVENTION: Compositions and Methods for  
TITLE OF INVENTION: Polynucleotide Delivery  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/620.925  
FILING DATE: 21-Jul-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/910.647  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-620-925-1

Query Match 3.2%; Score 49.4; DB 4; Length 9600;  
Best Local Similarity 49.8%; Pred. No. 0.0033;  
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGGCCCTGCAGGCCCACTGCGACGCTCATGAAGCAATATTACTGCTGGAGACCG 586  
DB 788 GCGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 847  
QY 587 ACAGTGCAGCCCTGCGCCAAAGGGGACCAAGTCTGAACTTCTGTATGACAGAGACCTG 646  
DB 848 GAGGGGCGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 907  
QY 647 ATTGATGACAGTAACTCAAGACATGCGAGTGAAGTCTTGAACATGAGCTGATGCTG 706  
DB 908 GAGGAGGGGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 967  
QY 707 GCTGACCGGCGCAGAGGAGAGATTGCGGACGCTGGAACAGGAACGTGATAAGTGGCAG 766  
DB 968 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1027  
QY 767 GTGCTAGTGA 777  
DB 1028 GAGGGGCGACGA 1038

RESULT 7

US-07-884-811-15

; Sequence 15, Application US/07884811

; Patent No. 5316921

; GENERAL INFORMATION:

; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.

; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/884,811

; FILING DATE: 19920518

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Dreger, Ginger R.

; REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: 755.1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-3216

; TELEFAX: 415/952-9881

; TELEK: 910/371-7168

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10596 bases

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

US-07-884-811-15

Query Match 3.2%; Score 49.4; DB 1; Length 10596;  
Best Local Similarity 49.8%; Pred. No. 0.0035;  
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGGCCCTGCAGGCCCACTGCGACGCTCATGAAGCAATATTACTGCTGGAGACCG 586  
DB 2322 GCGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2381  
QY 587 ACAGTGCAGCCCTGCGCCAAAGGGGACCAAGTCTGAACTTCTGTATGACAGAGACCTG 646  
DB 2382 GAGGGGCGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2441  
QY 647 ATTGATGACAGTAACTCAAGACATGCGAGTGAAGTCTTGAACATGAGCTGATGCTG 706  
DB 2442 GAGGAGGGGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2501  
QY 707 GCTGACCGGCGCAGAGGAGAGATTGCGGACGCTGGAACAGGAACGTGATAAGTGGCAG 766  
DB 2502 GGGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2561  
QY 767 GTGCTAGTGA 777  
DB 2562 GAGGGGCGACGA 2572

RESULT 8

US-07-885-971-15

; Sequence 15, Application US/07885971

; Patent No. 5328837

; GENERAL INFORMATION:

; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.

; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/885,971

; FILING DATE: 19920518

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Dreger, Ginger R.

; REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: 779

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-3216

; TELEFAX: 415/952-9881

; TELEK: 910/371-7168

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10596 bases

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

US-07-885-971-15

Query Match 3.2%; Score 49.4; DB 1; Length 10596;  
Best Local Similarity 49.8%; Pred. No. 0.0035;  
Matches 125; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 527 GCGAAGGCCCTGCAGGCCCACTGCGACGCTCATGAAGCAATATTACTGCTGGAGACCG 586  
DB 2322 GCGCAGAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2381





Qy	707	GCTACCCGGCCGCCGAAGCCAGAGATTCCGACCTGTAACAGAACTGCATTAAGTGGCAG	766
Db	2502	GGCCAGAGAGGGGCCAGAGCCAGAGAGGGCCAGAGAGGGCCAGAGAGCAGCAGAG	2561
Qy	767	GTGCTAATGGA	777
Db	2562	GAGGGCCAGGA	2572

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	Similarity	Score	DB 2	Length
Best Local	Similarity 49.8%	Pred. No. 0.0035		10596
Matches 125	Conservative 0	Mismatches 126	Indels 0	Gaps 0
QY	527	GGGAGAGCCCTGCAGGCCCATCTGGCAGCTCATTGACAGATTATTCCTGCTGGAGACCG	586	
Db	2322	GGGCGAGAGAGGGGCGAGGACGAGGAGGGGGGACGAGAGAGGAGGGGCGAGGCGAG	2381	
QY	587	AACAGTCAGACCCCTGCTCCCAAGGGGACCAAGTCTGAATCTTCTGTATCAGAGGACCG	646	
Db	2382	GAGGGGCGACGAGGAGGAGGAGGGGCGAGACGACGAGAGAGGGGCGAGGCGAGGACG	2441	
QY	647	ATTGATGACAGTAAGTCTCAAGACATCGAGATGAGATCCTGGAAATGAGCTGATGATG	706	
Db	2442	GAGGAGGGGCGACGAGAGGGGCGAGGAGGGGCGAGACGAGAGGGGCGAGGCGAGGAG	2501	
QY	707	GCTGACCGGCGCCAGAAAGCAGAGATTCGGCAGCTTGGAACAGAAATCTGATATAGTGGCAG	766	
Db	2502	GGGCGAGAGAGGCGCAGGACGAGAGGAGGGGCGACGAGAGGGGCGAGGCGGCGACGAGCAGGAG	2561	

```

CY      767 GTGCTAGTGA 777
|       |
Db      2562 GAGGGCGACGA 2572

RESULT 12
PCT-US93-04648-15
; Sequence 15 Application PC/TUS9304648
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Godowski, Paul J., Lokker, Natalie A., Mark, Melanie R
; TITLE OF INVENTION: HEPATOCTYE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/04648
; FILING DATE: 19930517
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755, 779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US93-04648-15

```

	Query Match	3.2%	Score 49.4	DB 5	Length 10596
	Best Local Similarity	49.8%	Pred. No. 0.0035		
	Matches 125	Conservative	0	Mismatches 126	Indels 0
				Gaps	0
QY	527	GGGAAGGCCCTTCAGAGCCCATGCTGACGCTCATGAAGCAGTATTACCTCTGTGAGAGACCAG	586		
DB	2322	GGGCGAGAGGGGGCGAGAGCAGAGAGAGGGGCGAGAGCAGAGAGAGGGGCGAG	2381		
QY	587	ACAGTGCAGCCGCTGCCCAAGGGGACCAAGTGTGAATTCTTGATGACAGGACCTG	646		
DB	2382	GAGGGGCGAGAGCAGAGAGAGGGGCGAGGACGAGAGAGGGGCGAGGGGCGAGGACAG	2441		
QY	647	ATTGATGACAGTAACTCAAGGACATGCGAGATGAGGTTCTTGGAACTATGACGTGATGGTG	706		
DB	2442	GAGGAGGGGCGAGGAGAGGGGCGAGGGGCGAGGCGAGGAGAGGGGCGAGGAGGAG	2501		
QY	707	GCTGACCGGCGCCAGAAACCGAGAGATTGGCGAGCTTGGAACCGGAACGTGATTAGTGGCAG	766		
DB	2502	GGGCGAGAGGGGCGAGAGCAGAGAGGAGGGGCGAGGAGGGCGAGGAGGCGAGGAGGAG	2561		
QY	767	GTGCTAGTGGGA	777		
DB	2562	GAGGGGCGAGGA	2572		

## RESULT 13

US-08-728-323A-1

Sequence 1, Application US/08728323A

Patent No. 5948676

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelmann, Isidore S.

TITLE OF INVENTION: Immediate Early Protein From Kaposi's

TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

TITLE OF INVENTION: Encoding Same And Uses Thereof

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper &amp; Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/728,323A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3489 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..3489

US-08-728-323A-1

Query Match 3.1%; Score 48.6; DB 2; Length 3489;

Best Local Similarity 49.8%; Pred. No. 0.0033;

Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```
OY 531 AGCCCTGAGGCCCACTGCGAGCTCATGAAGCAGATTACTGCTGAGGACGACAG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1901 AGCAGCAGCAGAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 1960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 591 TGCAGCCGCTGCCCAAGGGGACCAAGTGTGAACCTTCTCTGATCAGAGAGACTGATTG 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1961 AGCAGCAGCAGAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 2020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 651 ATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGAACATGACTGATGTGGCTG 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2021 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 711 ACCGGCGCCAGGAAGGAGAGATTTGGCAGCTGGAACAGAACTGATAGTGGCAGGTGC 770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2081 AGCAGCAGCAGATGACGACGACGAGATGACGAGATGACGAGAGATGACGAGAGAC 2140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 771 TAGTGA 777
    |||
```

DB 2141 AGCAGGA 2147

## RESULT 14

US-09-298-568-1

Sequence 1, Application US/09298568

Patent No. 6322792

GENERAL INFORMATION:

APPLICANT: Kieff, Elliott D.

APPLICANT: Ballestas, Mary E.

TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO

TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE

FILE REFERENCE: 16412-1001R

CURRENT APPLICATION NUMBER: US/09/298,568

CURRENT FILING DATE: 1998-04-21

EARLIER APPLICATION NUMBER: US 60/109,422

EARLIER FILING DATE: 1998-11-19

NUMBER OF SEQ ID NOS: 3

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 1

LENGTH: 3489

TYPE: DNA

ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-298-568-1

Query Match 3.1%; Score 48.6; DB 4; Length 3489;

Best Local Similarity 49.8%; Pred. No. 0.0033;

Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```
OY 531 AGCCCTGAGGCCCACTGCGAGCTCATGAAGCAGATTACTGCTGAGGACGACAG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1901 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 1960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 591 TGCAGCCGCTGCCCAAGGGGACCAAGTGTGAACCTTCTCTGATCAGAGAGACTGATTG 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1961 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 2020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 651 ATGACAGTAAGCTCAAGACATGCGAGATGAGTCTCTGGAACATGACTGATGTGGCTG 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2021 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 711 ACCGGCGCCAGGAAGGAGAGATTTGGCAGCTGGAACAGAACTGATAGTGGCAGGTGC 770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2081 AGCAGCAGCAGATGACGACGACGAGATGACGAGATGACGAGAGATGACGAGAGAC 2140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 771 TAGTGA 777
    |||
DB 2141 AGCAGGA 2147
```

```
RESULT 15
US-08-770-379-20/c
Sequence 20, Application US/08770379
Patent No. 5849564
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelmann, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/770,379  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Query Match 3.1%; Score 48.6; DB 2; Length 32207;  
Best Local Similarity 49.8%; Pred. No. 0.0099;  
Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
QY 531 AGGCCCTGCAGGCCCACTGGCAGCTCATGAAGCAGTATTACCTGCTGAGAGACCAGACAG 590  
DB 20096 AGCAGCAGCAGAGTAGAGCAGCAGCAGAGATGAGCAGCAGATGAGCAGCAGAGATG 20037  
QY 591 TGCAGCCGCTGCCCAAGGGGACCAAGTGTGTAATTCTTCTGATGACAGAGACCTGATTG 650  
DB 20036 AGCAGCAGCAGAGTAGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGT 19977  
QY 651 ATGACAGTAAGCTCAAGGACATGCGAGATGAGGTCTGGAACTAGCTGATGATGGCTG 710  
DB 19976 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGAGATG 19917  
QY 711 ACCGGGCGCCAGAAAGCGAGATTGGCAGCTGGAAACAGAACTGCAATGAAGTGGCAGGTGC 770  
DB 19916 AGCAGCAGCAGATGAGCAGAGAGCAGCAGAGATGAGCAGAGAGCAGAGATGAGCAGAGC 19857  
QY 771 TAGTGGA 777  
DB 19856 AGCAGGA 19850

Search completed: March 27, 2003, 07:50:43  
Job time : 223 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 27, 2003, 07:46:17 ; Search time 186 Seconds  
(without alignments)  
7105.167 Million cell updates/sec

Title: US-09-744-125a-1  
Perfect score: 1553  
Sequence: 1 gaatcgcgcagcaggtcg99.....aaaaaaaaaacctcgag 1553

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCr\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCrUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	214.6	13.8	371 10	US-09-783-590-2176 Sequence 2176, Ap
2	157.8	10.2	318 10	US-09-783-590-10575 Sequence 10575, A
3	101.2	6.5	385 10	US-09-783-590-10146 Sequence 10146, A
4	71.6	4.7	457 10	US-09-764-887-30 Sequence 30, Appl
5	58	3.6	2108 10	US-09-962-832-225 Sequence 225, App
6	46.6	3.0	8459 10	US-09-817-913-8 Sequence 8, Appl
7	46.6	3.0	8459 10	US-09-817-913-8 Sequence 8, Appl
8	45.4	2.9	3131 10	US-09-817-913-14 Sequence 14, Appl
9	45.4	2.9	3131 10	US-09-817-913-14 Sequence 14, Appl
10	45.2	2.9	14800 10	US-09-954-856-1601 Sequence 1601, Ap
11	43.4	2.8	7305 9	US-10-156-239-9 Sequence 9, Appl
12	43.4	2.8	7305 9	US-09-795-693-9 Sequence 9, Appl
13	43.4	2.8	8056 9	US-10-072-621-3 Sequence 3, Appl
14	43.4	2.8	8195 10	US-10-156-239-7 Sequence 7, Appl
15	43.4	2.8	8195 10	US-09-785-693-7 Sequence 7, Appl
16	42.4	2.7	2553 10	US-09-815-242-7690 Sequence 7690, Ap
17	42.2	2.7	422 10	US-09-854-133-337 Sequence 337, App
18	42.2	2.7	422 10	US-09-738-973-337 Sequence 337, App
19	42	2.7	765 9	US-09-925-299-143 Sequence 143, App

20	42	2.7	765 10	US-09-925-299-143 Sequence 143, App
21	42	2.7	1852 10	US-09-969-852-4 Sequence 24, Appl
22	40.6	2.6	485 9	US-09-992-598-245 Sequence 245, App
23	40.6	2.6	485 9	US-09-989-293A-245 Sequence 245, App
24	40.6	2.6	485 9	US-10-063-547-43 Sequence 43, Appl
25	40.6	2.6	485 9	US-09-989-735-245 Sequence 245, App
26	40.6	2.6	485 9	US-09-990-444-245 Sequence 245, App
27	40.6	2.6	485 9	US-09-989-730-245 Sequence 245, App
28	40.6	2.6	485 9	US-09-990-436-245 Sequence 245, App
29	40.6	2.6	485 9	US-09-991-181-245 Sequence 245, App
30	40.6	2.6	485 9	US-09-993-687-245 Sequence 245, App
31	40.6	2.6	485 9	US-09-989-734-245 Sequence 245, App
32	40.6	2.6	485 9	US-09-997-653-245 Sequence 245, App
33	40.6	2.6	485 9	US-10-174-590-185 Sequence 185, App
34	40.6	2.6	485 9	US-10-176-758-185 Sequence 185, App
35	40.6	2.6	485 9	US-10-063-616-43 Sequence 43, Appl
36	40.6	2.6	485 9	US-10-175-737-185 Sequence 185, App
37	40.6	2.6	485 9	US-09-993-667-245 Sequence 245, App
38	40.6	2.6	485 9	US-10-063-502-43 Sequence 43, Appl
39	40.6	2.6	485 9	US-10-173-706-185 Sequence 185, App
40	40.6	2.6	485 9	US-10-175-738-185 Sequence 185, App
41	40.6	2.6	485 9	US-10-176-482-185 Sequence 185, App
42	40.6	2.6	485 9	US-10-176-482-185 Sequence 185, App
43	40.6	2.6	485 9	US-10-176-757-185 Sequence 185, App
44	40.6	2.6	485 9	US-10-176-913-185 Sequence 185, App
45	40.6	2.6	485 9	US-10-180-552-185 Sequence 185, App

#### ALIGNMENTS

RESULT 1  
US-09-783-590-2176  
Sequence 2176, Application US/09783590  
Patent No. US20020110850A1  
GENERAL INFORMATION:  
APPLICANT: Dillion, Patrick J.  
APPLICANT: Haselton, William A.  
APPLICANT: Li, Haodong  
APPLICANT: Rosen, Craig A.  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2  
FILE REFERENCE: PO-16.2C1  
CURRENT APPLICATION NUMBER: US/09/783,590  
CURRENT FILING DATE: 2000-02-15  
PRIOR APPLICATION NUMBER: 08/420,856  
PRIOR FILING DATE: 1995-04-12  
PRIOR APPLICATION NUMBER: 08/346,731  
PRIOR FILING DATE: 1994-11-21  
NUMBER OF SEQ ID NOS: 12485  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2176  
LENGTH: 371  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (106)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (114)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (117)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (125)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (131)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (146)



```
RESULT 3
US-09-783-590-10146
; Sequence 10146, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseldine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10146
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (237)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (309)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (340)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (352)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (367)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (379)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (385)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-10146

Query Match          6.5%; Score 101.2; DB 10; Length 385;
Best Local Similarity 95.1%; Pred. No. 7.9e-21;
Matches 136; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
```

```
; Sequence 30, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult P41M or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (387)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-887-30
```

```
Query Match          4.6%; Score 71.6; DB 10; Length 457;
Best Local Similarity 94.4%; Pred. No. 1e-11;
Matches 85; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Oy 958 AGGTGTCATCAAGCTGAAGACAACAGGATTTCTTCATGCGCATGAGGTCGACGGCC 1017
Db 276 AGGTGTCATCAAGCTGAAGACAACAGGATTTCTTCATGCGCATGAGGTCGACGG-C 334

Oy 1018 CATCTACATGATGAGCGCGCGGTGCTGTG 1047
Db 335 CATCTACATGATGAGCGCGGTGCTGTG 364
```

```
RESULT 5
US-09-962-832-225
; Sequence 225, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-225
```

```
Query Match          3.7%; Score 58; DB 10; Length 2108;
Best Local Similarity 53.0%; Pred. No. 3.3e-07;
Matches 124; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Oy 536 CTGACGGCCCACTGGAGCTGATGAAGACATTAACCTGCTGAGAGACAGACAGTGGAG 595
Db 1056 CTGAGACACTGGAGACAGAGAGGGGAGCTGGAGCACTGGAGACAGACAGAGAGGGCAG 1115

Oy 596 CCGTGCCCAAGGGGAGCAAGATGCTGAATCTCTGTATGACAGAGACCTGATTGATAC 655
Db 1116 CTGGGGCTCCAGAGACAGACAGATGCTGACGCTGAAGACGCTTGAAGAGACAGACAGGGGAG 1175

Oy 656 AGTAAGCTCAAGACATGCGAGATGAGGTCTTGGAACATGAGCTGATGTGCTGACCGG 715
Db 1176 CCAAGCACTGAGAGAGAGAGGGGAGGCTGAAGCACTGTGTGCACAGAGAGGGGAG 1235
```

QY 716 GCGCAGAGCGAGATTGCGAGCTGGAACAGAACTGCTAATAGTGGCAGG 769  
DB 1236 CTGAAGCATCTGTGACAGAGAGGGGCGCTGAGACAGACAGAGGAGGAGT 1289

## RESULT 6

US-09-817-913-8  
; Sequence 8, Application US/09817913  
; Patent No. US2002061860A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Zuomei  
; APPLICANT: Bonfils, Claire  
; APPLICANT: Besterman, Jeffrey  
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms  
; FILE REFERENCE: 106101.145  
; CURRENT APPLICATION NUMBER: US/09/817,913  
; PRIOR FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/192,157  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 8459  
; TYPE: DNA  
; ORGANISM: Human  
US-09-817-913-8

Query Match 3.0%; Score 46.6; DB 10; Length 8459;  
Best Local Similarity 49.8%; Pred. No. 0.0022;  
Matches 118; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 532 GGGCTGCGAGGCCCACTGCGAGCTCATGAAGCATTTACTGCTGGAGAGCAACAGCACT 591  
DB 1020 GGGCTCAAGCAGAAACACAGATCCAGAGCATCTCATGCGGAGTTCCAGAGGCA 1079  
QY 592 GCAGCGCGCTGCCAAAGGGGACCAAGTGTGAATTCTCTGATGACAGAGCACTGATTGA 651  
DB 1080 GCACGAGCAGCTCTCCCGGACAGACAGGCGCAGCTCCAGAGACATCAAGCAATTAACA 1139  
QY 652 TGACAGTAAGCTCAAGAGCATGAGATGAGTCTTGAAGCATGAGCTGATGAGGCTGAG 711  
DB 1140 GGAGATGCTGGCATGAACACAGCAGAGAGCTGTGGAACACAGCGAAGCTGGAGAG 1199  
QY 712 CCGGGGCCAGAGCGAGATTCGCGAGCTGGAACAGAACTGCTAATAGTGGCAGGT 768  
DB 1200 GCACCGCCAGAGCAGAGAGCTGGAAGACAGCACCGGAGAGCAAGACTGACAGAGCT 1256

## RESULT 7

US-09-817-538-8  
; Sequence 8, Application US/09817538  
; Patent No. US20020137162A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Zuomei  
; APPLICANT: Bonfils, Claire  
; APPLICANT: Besterman, Jeffrey  
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone  
; FILE REFERENCE: 106101.144  
; CURRENT APPLICATION NUMBER: US/09/817,538  
; PRIOR FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/192,157  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 8459  
; TYPE: DNA  
; ORGANISM: Human  
US-09-817-538-8

Query Match 3.0%; Score 46.6; DB 10; Length 8459;  
Best Local Similarity 49.8%; Pred. No. 0.0022;

Matches 118; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 532 GGGCTGCGAGGCCCACTGCGAGCTCATGAAGCATTTACTGCTGGAGAGCAACAGCACT 591  
DB 1020 GGGCTCAAGCAGAAACAGAGCATCCAGAGCATCTCATGCGGAGTTCCAGAGGCA 1079  
QY 592 GCAGCGCGCTGCCAAAGGGGACCAAGTGTGAATTCTCTGATGACAGAGCACTGATTGA 651  
DB 1080 GCACGAGCAGCTCTCCCGGACAGACAGGCGCAGCTCCAGAGACATCAAGCAATTAACA 1139  
QY 652 TGACAGTAAGCTCAAGAGCATGAGATGAGTCTTGAAGCATGAGCTGATGAGTGGCTGA 711  
DB 1140 GGAGATGCTGGCATGAACACAGCAGAGAGCTGTGGAACACAGCGAAGCTGGAGAG 1199  
QY 712 CCGGGGCCAGAGCGAGATTCGCGAGCTGGAACAGAACTGCTAATAGTGGCAGGT 768  
DB 1200 GCACCGCCAGAGCAGAGAGCTGGAAGACAGCACCGGAGAGCAAGACTGACAGAGCT 1256

## RESULT 8

US-09-817-913-14  
; Sequence 14, Application US/09817913  
; Patent No. US2002061860A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Zuomei  
; APPLICANT: Bonfils, Claire  
; APPLICANT: Besterman, Jeffrey  
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms  
; FILE REFERENCE: 106101.145  
; CURRENT APPLICATION NUMBER: US/09/817,913  
; PRIOR FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/192,157  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 3131  
; TYPE: DNA  
; ORGANISM: Human  
US-09-817-913-14

Query Match 2.9%; Score 45.4; DB 10; Length 3131;  
Best Local Similarity 51.8%; Pred. No. 0.003;  
Matches 103; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 712 CCGGGGCCAGAGCGAGATTCGCGAGCTGGAACAGAACTGCTAATAGTGGCAGGTCT 771  
DB 1796 CCGAGGCGGAGAGGCTCCCTGGAAGAGCTGCAGTGCTCACTGAGCGGCACTGCT 1855  
QY 772 AGTGAACAGATCAACAGCATGAGCTCCGAGCTTTCGACAAACAGACACTGGCAGTGT 831  
DB 1856 CCTTACGCGACAAACCCGCTAGCGCTCAACTGGAACAGGAGAGCTGGCAGGCT 1915  
QY 832 GCGGGCCGATGTCGCTGCTGATGAGTGGCTGCGTGAATCACTGGGAGAGCAAC 891  
DB 1916 CTTGCAACAGCGAGATTTGAGATGCTGCTGTGTGGGTGGGGTGGAGACAGAC 1975  
QY 892 CAAGATTAACAGATTGAT 910  
DB 1976 CATCTGAATGAGCTTCAT 1994

## RESULT 9

US-09-817-538-14  
; Sequence 14, Application US/09817538  
; Patent No. US20020137162A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Zuomei  
; APPLICANT: Bonfils, Claire  
; APPLICANT: Besterman, Jeffrey  
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone  
; FILE REFERENCE: 106101.144



;; CURRENT APPLICATION NUMBER: US/09/817,538  
;; CURRENT FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: US 60/192,157  
;; PRIOR FILING DATE: 2000-03-24  
;; NUMBER OF SEQ ID NOS: 33  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 14  
;; LENGTH: 3131  
;; TYPE: DNA  
;; ORGANISM: Human  
US-09-817-538-14

Query Match 2.9%; Score 45.4; DB 10; Length 3131;  
Best Local Similarity 51.8%; Pred. No. 0.003;  
Matches 103; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 712 CCGGCGCCAGAGCGAGATTCCGCGAGCTGGAACAGAACTGATTAAGTGGCAGTGTCT 771  
DB 1796 CCGAGGCGCGAGGCTCTCCCTGGAAGAGCTGAGTCCACTCTGAGCGGCGACGTGCT 1855  
QY 772 AGTGAGACAGATCAAGGAGATGCTCTCCGAGCTTGACACAGACACTGGGAGTGTCT 831  
DB 1856 CCTTACGAGCACCAACCCGCTCAGCCGCTCAAACTGACACAGGAGAGCTGGGAGGCT 1915  
QY 832 GCGGCGCCGAGTGGTGGGATCTGATGCGCTCCGCTGAGATCAACCTGGGCGACAGCAAC 891  
DB 1916 CTTGGCAGGAGGATGTTTGAGATGCTGCTGCTGGTGGGCTGGGAGTGGACACTGACAC 1975  
QY 892 CAAGATTAACCAATGAT 910  
DB 1976 CATCTGGAATGAGCTTCAT 1994

RESULT 10  
US-09-954-456-1601  
;; Sequence 1601, Application US/09954456  
;; Patent No. US20020115057A1  
;; GENERAL INFORMATION:

;; APPLICANT: Young, Paul  
;; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
;; FILE REFERENCE: 689290-76  
;; CURRENT APPLICATION NUMBER: US/09/954,456  
;; CURRENT FILING DATE: 2001-09-18  
;; PRIOR APPLICATION NUMBER: US/60/233,617  
;; PRIOR FILING DATE: 2000-09-18  
;; PRIOR APPLICATION NUMBER: US/60/234,052  
;; PRIOR FILING DATE: 2000-09-20  
;; PRIOR APPLICATION NUMBER: US/60/234,923  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: US/60/235,134  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: US/60/235,637  
;; PRIOR FILING DATE: 2000-09-26  
;; PRIOR APPLICATION NUMBER: US/60/235,638  
;; PRIOR FILING DATE: 2000-09-26  
;; PRIOR APPLICATION NUMBER: US/60/235,711  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: US/60/235,720  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: US/60/235,840  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: US/60/235,863  
;; PRIOR FILING DATE: 2000-09-27  
;; NUMBER OF SEQ ID NOS: 2276  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 1601  
;; LENGTH: 14800  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-954-456-1601

Query Match 2.9%; Score 45.2; DB 10; Length 14800;

Best Local Similarity 46.1%; Pred. No. 0.008;  
Matches 189; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

QY 360 TGAGGAGCTGACACCCAGAGGCTATTTCAGCCATCCAGACCAAGGCCCTGTTAGCAAG 419  
DB 7535 TGCTGAGCTCAAGTCTGAGAGATGCAACGGTGCAGAGACAGCTCTGACAGAGA 7594  
QY 420 CTGAGAGCAGCTGCTGAGAAAGT---GGGATGCAACCAAGCCCACTTGGAGACT 476  
DB 7595 CGGAGCCCTGCGACCAAGCTTCTCTGAAAAGACAGCTCTACAGGGAGCGCT 7654  
QY 477 TCCAGACTGCTGACAGACACCTGATGCTTCTGACCTGGCCGTAACGCGAAGGCC 536  
DB 7655 TCATGAGCAGAGAAAGCCAGAGCTGAGCAGCTTCTCAGAGAGAGGTGGCCAAAGCAC 7714  
QY 537 TGCAGGCCCACTGCGACCTATGAAAGATTAACCTGCTGAGAGACCAAGCACTGACAG 596  
DB 7715 AGCAGCTGCTGAGAGACAGACGCGGACAGCAGAGATGAGACAGGAACGCGAGCGGC 7774  
QY 597 CGCTGCCCAAGGGGACCAAGTGTGAATCTCTGATGACAGAGACCTGATTGATGACA 656  
DB 7775 TGTGTCGACGATGAGAGAGGCGCGCGCGGACGATGAGGCCGAGAGGGCTGCGGC 7834  
QY 657 GTAAGCTCAAGGACATGCGAGATGAGTCTTGAACATGAGCTGATGCTGACCGGC 716  
DB 7835 GCAAGCAGAGAGAGCTGACAGAGCTGAGACAGACGCGCGGACAGAGAGAGCTGTGG 7894  
QY 717 GCCAAGACGAGAGATTGCGCAGCTGGAACAGAACTGCAATTAAGTGGCAG 766  
DB 7895 CTGAGGAGAACCAAGAGCTGAGAGCAGCTGACAGCTCTGAGAGAGCAG 7944

RESULT 11  
US-10-156-239-9  
;; Sequence 9, Application US/10156239  
;; Publication No. US20030036074A1  
;; GENERAL INFORMATION:

;; APPLICANT: Glucksmann, Maria A.  
;; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Transl  
;; TITLE OF INVENTION: Arpase Molecule, A Human Ubiquitin Hydrolyase-Like Molecule, A Hu  
;; FILE REFERENCE: 35800/247645  
;; CURRENT APPLICATION NUMBER: US/10/156,239  
;; CURRENT FILING DATE: 2002-05-24  
;; PRIOR APPLICATION NUMBER: 09/795,693  
;; PRIOR FILING DATE: 2001-02-28  
;; PRIOR APPLICATION NUMBER: 60/185,906  
;; PRIOR FILING DATE: 2000-02-29  
;; PRIOR APPLICATION NUMBER: 09/809,557  
;; PRIOR FILING DATE: 2001-03-15  
;; PRIOR APPLICATION NUMBER: 60/192,018  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: 09/808,568  
;; PRIOR FILING DATE: 2001-03-14  
;; PRIOR APPLICATION NUMBER: 60/191,790  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: 09/808,767  
;; PRIOR FILING DATE: 2001-03-15  
;; PRIOR APPLICATION NUMBER: 60/191,781  
;; PRIOR FILING DATE: 2000-03-24  
;; NUMBER OF SEQ ID NOS: 60  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 9  
;; LENGTH: 7305  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-156-239-9

Query Match 2.9%; Score 43.4; DB 9; Length 7305;  
Best Local Similarity 45.3%; Pred. No. 0.019;  
Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 803 GACTTCGACAAACAGACACTGCGAGTCTGCGGGCCGCGATGGTGGTACCTGATGCGC 862  
Db 2902 GCCATGAGAGAGCGCGCTTTAGAGAGACCCGTGGCATGGAGAGAGGCCCAACCTG 2961  
QY 863 TCGCGTAGATACCTTGGGAGAGCAACCAAGATTAACAGATTGATGTAAGCTGTCT 922  
Db 2962 CCTGTGTTGTCTGCTGAGCAAACTCAACAAGTCTCAAGAGACGACAAAGAGCTGGCC 3021  
QY 923 CTGAGAGGCTCGGCTGGAAGATATCCCGGAAAACAAGTGTATCAAGCTGAAGAACAC 982  
Db 3022 CTGAACAAGCTAGGCTTAACCTTAAGAGAACAGAGTGTCTCTTCTTGGGCCCAAC 3081  
QY 983 GGTGATTTCTTATTTGATGCAATGAGGGTCAAGGCCCATTAATCATGATGAGAGCGCGGTG 1042  
Db 3082 GGGGGGGGCAACCAACCAACCAATGTCATCTGACCGGCTGTCTCCCTCAACGTGGGT 3141  
QY 1043 CTCTGTGCTTCCAAATGCGGCTCAGCAACAACTGTGTGAGATGCGCAGCTGCGA 1102  
Db 3142 TCCGCCACCATCTAGCGGAGACGACATCCGACGAGATGATGATGATCCGCAAGAACTG 3201  
QY 1103 TTGCTCTTCTTATCAACCAAGAGCTCATTTGCCCTTCATCAGGGCTGAGG 1151  
Db 3202 GGCATGTGCGCGAGACGACATGTGCTTTTGAACGCGCTCACGCTGAGG 3250

## RESULT 12

US-09-795-693-9  
; Sequence 9, Application US/09795693  
; Patent No. US20020068710A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria A.  
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and  
; TITLE OF INVENTION: 32613, No. US20020068710A1 Human Transporters  
; FILE REFERENCE: 35800/209292  
; CURRENT APPLICATION NUMBER: US/09/795,693  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/185,906  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 7305  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-795-693-9

Query Match 2.8%; Score 43.4; DB 10; Length 7305;

Best Local Similarity 45.3%; Pred. No. 0.019;

Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 803 GACTTCGACAAACAGACACTGCGAGTCTGCGGGCCGCGATGGTGGTACCTGATGCGC 862  
Db 2902 GCCATGAGAGAGCGCGCTTTAGAGAGACCCGTGGCATGGAGAGAGGCCCAACCTG 2961  
QY 863 TCGCGTAGATACCTTGGGAGAGCAACCAAGATTAACAGATTGATGTAAGCTGTCT 922  
Db 2962 CCTGTGTTGTCTGCTGAGCAAACTCAACAAGTCTCAAGAGACGACAAAGAGCTGGCC 3021  
QY 923 CTGAGAGGCTCGGCTGGAAGATATCCCGGAAAACAAGTGTATCAAGCTGAAGAACAC 982  
Db 3022 CTGAACAAGCTAGGCTTAACCTTAAGAGAACAGAGTGTCTCTTCTTGGGCCCAAC 3081  
QY 983 GGTGATTTCTTATTTGATGCAATGAGGGTCAAGGCCCATTAATCATGATGAGAGCGCGGTG 1042  
Db 3082 GGGGGGGGCAACCAACCAACCAATGTCATCTGACCGGCTGTCTCCCTCAACGTGGGT 3141  
QY 1043 CTCTGTGCTTCCAAATGCGGCTCAGCAACAACTGTGTGAGATGCGCAGCTGCGA 1102  
Db 3142 TCCGCCACCATCTAGCGGAGACGACATCCGACGAGATGATGATGATCCGCAAGAACTG 3201  
QY 1103 TTGCTCTTCTTATCAACCAAGAGCTCATTTGCCCTTCATCAGGGCTGAGG 1151  
Db 3202 GGCATGTGCGCGAGACGACATGTGCTTTTGAACGCGCTCACGCTGAGG 3250

## RESULT 13

US-10-072-621-3  
; Sequence 3, Application US/10072621  
; Patent No. US20020169137A1  
; GENERAL INFORMATION:  
; APPLICANT: Reiner, Peter B.  
; APPLICANT: Connop, Bruce P.  
; APPLICANT: Pollard, Michelle  
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION  
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY  
; FILE REFERENCE: 100103,402  
; CURRENT APPLICATION NUMBER: US/10/072,621  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 8056  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-072-621-3

Query Match 2.8%; Score 43.4; DB 9; Length 8056;

Best Local Similarity 45.3%; Pred. No. 0.02;

Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 803 GACTTCGACAAACAGACACTGCGAGTCTGCGGGCCGCGATGGTGGTACCTGATGCGC 862  
Db 2952 GCCATGAGAGAGCGCGCTTTAGAGAGACCCGTGGCATGGAGAGAGGCCCAACCTG 3011  
QY 863 TCGCGTAGATACCTTGGGAGAGCAACCAAGATTAACAGATTGATGTAAGCTGTCT 922  
Db 3012 CCTGTGTTGTCTGCTGAGCAAACTCAACAAGTCTTCAAGAGAGCAAGAGCTGGCC 3071  
QY 923 CTGAGAGGCTCGGCTGGAAGATATCCCGGAAAACAAGTGTATCAAGCTGAAGAACAC 982  
Db 3072 CTGAACAAGCTAGGCTTAACCTTAAGAGAACAGAGTGTCTCTTCTTGGGCCCAAC 3131  
QY 983 GGTGATTTCTTATTTGATGCAATGAGGGTCAAGGCCCATTAATCATGATGAGAGCGCGGTG 1042  
Db 3132 GGGGGGGGCAACCAACCAACCAATGTCATCTGACCGGCTGTCTCCCTCAACGTGGGT 3191  
QY 1043 CTCTGTGCTTCCAAATGCGGCTCAGCAACAACTGTGTGAGATGCGCAGCTGCGA 1102  
Db 3192 TCCGCCACCATCTAGCGGAGACGACATCCGACGAGATGATGATGATCCGCAAGAACTG 3251  
QY 1103 TTGCTCTTCTTATCAACCAAGAGCTCATTTGCCCTTCATCAGGGCTGAGG 1151  
Db 3252 GGCATGTGCGCGAGACGACATGTGCTTTTGAACGCGCTCACGCTGAGG 3300

## RESULT 14

US-10-156-239-7  
; Sequence 7, Application US/10156239  
; Publication No. US20030036074A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: No. US20030036074A1 Nucleic Acid Sequences Encoding Human Trans  
; TITLE OF INVENTION: Atpase Molecule, A Human Ubiquitin Hydrolyase-Like Molecule, A Hu  
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor  
; FILE REFERENCE: 35800/247645  
; CURRENT APPLICATION NUMBER: US/10/156,239  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 09/795,693  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/185,906  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: 09/809,557  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/192,018  
; PRIOR FILING DATE: 2000-03-24

```

? PRIOR APPLICATION NUMBER: 09/808,568
? PRIOR FILING DATE: 2001-03-14
? PRIOR APPLICATION NUMBER: 60/191,790
? PRIOR FILING DATE: 2000-03-24
? PRIOR APPLICATION NUMBER: 09/808,767
? PRIOR FILING DATE: 2001-03-15
? PRIOR APPLICATION NUMBER: 60/191,781
? PRIOR FILING DATE: 2000-03-24
? NUMBER OF SEQ ID NOS: 60
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 7
? LENGTH: 8195
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (132)...(7442)
? OS-10-156-239-7

```

Query Match	2.8%;	Score 43.4;	DB 9;	Length 8195;
Best Local Similarity	45.3%;	Pred. No. 0.021;		
Matches 158; Conservative	0;	Mismatches 191;	Indels 0;	Gaps 0;

Qy	803	GATTGCACACACAGACATCTGGAGGCTGCGGGGCGCATGGTGGGATCTGATGGC	862
Db	3033	GCCATGGAGAGCCGGCGCTTTTGGAGAGACCCGTGGATGGAGAGAGCCACCCACTG	3092
Qy	863	TCGGGTGAGATCACTCTGGGCGAGAGACCAAGATTAACAGATTGATGGACCTGTCT	922
Db	3093	CCTTGGTTGTCTGCGTGGACAAATCTACCCAGGTCTACAAAGACGACAAAGAGTGGCC	3152
Qy	923	CTGAGGGGTCCGGCCTTGGAGATATCCCGAAAACAAGTGTATCAAGCTGAAGAACAC	982
Db	3153	CTGAACAAGGTGAGCTGAACCTCTAGAGAAACAAGTGTCTCCTTCTTGGCCACACAC	3212
Qy	983	GGTGATTTCTTTCATTGCAATAGAGGTCGACGGCCATTTACATGATGAGACGCGGTG	1042
Db	3213	GGGGCGGGCAAGACCAACCATGATCCATCTGACCGGCTGTTCCTTCCAACGTGGGT	3272
Qy	1043	CTCTGTGGCTTCAAAATGGCGGCTCAGCAACACTCTGTGTGGAAATCGCCAGCTGGGA	1102
Db	3273	TCCGCCACCATCTACGGGACGACATCCGACGGAGATGAATGCAATCGCAAGAACTG	3322
Qy	1103	TTTGCTCTCTTCAACCAAGGACCTCATGGCCCTCATGAGGCTGAGG	1151
Db	3333	GGCATGTGCCGACGACACATGTGTCTTTTGAACCGGCTACGCTGAGG	3381

```

RESULT 15
US-09-795-693-7
Sequence 7, Application US/09795693
Patent No. US20020068710A1
GENERAL INFORMATION:
APPLICANT: Gluckemann, Maria A.
TITLE OF INVENTION: 20665, 579, 17114, 23821, 33894, and
FILE OF INVENTION: 32613, No. US20020068710A1el Human Transporters
FILE REFERENCE: 35800/209292
CURRENT APPLICATION NUMBER: US/09/795,693
CURRENT FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/185,906
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 42
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 8195
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (132)...(7442)
US-09-795-693-7

```

```

Best Local Similarity: 45.3%; Pred. No. 0.021;
Matches 158; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 803 GACTTCGACCAACGAGACCTGGCAGTCTGCGGGGCGGCATGTGCGGTACTGTATGCGC 862
Db 3033 GGCATGGAGAGCCGGCCGCTTTGAGAGAGCCCGTGGCATGTAGGAGGAGCCACCACTG 3092

QY 863 TCGCGTGAATACCTCTGGGAGAGACAACAAAGATTAACAGATGATGTGACTGTCT 922
Db 3093 CCTCGTGTGTCTCGGTGAGACAAACTCAACAAAGGTCTTCAAGAGACGAACAAGACGTGGCC 3152

QY 923 CTGAGGGTCCGGCCTGGAAGATATCCCGGAAACAAGTGTATCAACTGAAGAAACAAC 982
Db 3153 CTGAACAAGCTGAGGCTGTAACTCTACAGAGAACCAAGGTGTCTCTTCTTGGGCCACAAC 3212

QY 983 GGTGATTTCTTCATTTGCGCAATGAGGGGTGACGAGCCATCTACATGATGAGACGCGCGTG 1042
Db 3213 GGGGCGGGGACGAACACCAACCATGTTCATCTGTGACCGGCTGTTCCTTCCAAAGTCGGGT 3272

QY 1043 CTCTGTGGCTCCAAATGAGCGGCTCAGCAACAATCTGTGTGTGAGAGATGCCAGCCCTGCGA 1102
Db 3273 TCCGCGACACATCTACGCGGACGAGATCCGACAGGAGATGATGAGATCCGCAAGAACCTG 3332

QY 1103 TTTCGTTCTCTTATCAACCAAGACTCTATTGCTCCTCATCAGGGCTGAGG 1151
Db 3333 GGCATGTGTCGCGAGACACAATGTGCTTTTGAACCGGCTACAGGTGAGG 3381

Search completed: March 27, 2003, 09:16:22
Job time : 272 secs

```

Query Match	2.8%	Score 43.4	DB 10	Length 8195
-------------	------	------------	-------	-------------

This Page Blank (uspto)



enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : [filang@lifetech.com](mailto:filang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 180 a 275 c 273 g 227 t 11 others  
ORIGIN

Query Match 59.5%; Score 924.2; DB 9; Length 966;  
Best Local Similarity 97.8%; Pred. No. 1.9e-162;  
Matches 945; Conservative 10; Mismatches 9; Indels 2; Gaps 2;

QY 448 ATGACACGACGAGCCACCTTGAGACCTTCCAGACCTGTCACAGACACCCCTGATGC 507  
DB 966 ATGACACGACGAGCCACCTTGAGACCTTCCAGACCTGTCACAGACACCCCTGATGC 507  
QY 508 CTCTTACCTGCGCCGTTACCGCGAAGCCCTGCAAGGCCCTGACCTGACCTATGAAGCACTA 567  
DB 906 CTCTTACCTGCGCCGTTACCGCGAAGCCCTGCAAGGCCCTGACCTGACCTATGAAGCACTA 567  
QY 568 TTACCGTGGAGGAGCAGACAGTGCAGCGCGCCCAAGGGGACCAAGTGTGAACCTT 627  
DB 846 TTACCGTGGAGGAGCAGACAGTGCAGCGCGCCCAAGGGGACCAAGTGTGAACCTT 627  
QY 628 CTCTGATGACAGAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687  
DB 786 CTCTGATGACAGAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727  
QY 688 GGAACATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747  
DB 726 GGAACATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747  
QY 748 GGAACATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807  
DB 666 GGAACATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807  
QY 808 CGACAAAC-CAGACACTGAGAGTGCCTGCGGGGCGGAGTGTGCGGTACCTGATGCGCTGCG 866  
DB 606 CGACAAACACTGAGAGTGCCTGCGGGGCGGAGTGTGCGGTACCTGATGCGCTGCG 866  
QY 867 GTGAGATGACCTGCGGAGAGCAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC 925  
DB 546 GTGAGATGACCTGCGGAGAGCAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC 925  
QY 926 GAGGGGCGGAGAGATATCCCGGAAACAGAGTGTATCAAGCTGAAGAACACAGG 985  
DB 486 GAGGGGCGGAGAGATATCCCGGAAACAGAGTGTATCAAGCTGAAGAACACAGG 985  
QY 986 GATTTCTTATGTCATGAGAGGTGACGCGCCATCTACATGATGAGAGCGCGGAGCTC 1045  
DB 426 GATTTCTTATGTCATGAGAGGTGACGCGCCATCTACATGATGAGAGCGCGGAGCTC 1045  
QY 1046 TGTGGCTCCAATGCGGCTCAGCAACAACTCTGTGTGTGAGATGCGCAGAGCTGCGATTC 1105  
DB 366 TGTGGCTCCAATGCGGCTCAGCAACAACTCTGTGTGTGAGATGCGCAGAGCTGCGATTC 1105  
QY 1106 GTCTTCTTATCAACAGAGACCTCATTTGCTCATGAGGGCTGAGGCTGCAAGATCA 1165  
DB 306 GTCTTCTTATCAACAGAGACCTCATTTGCTCATGAGGGCTGAGGCTGCAAGATCA 1165  
QY 1166 CCACAGTGAAGATGAGTGGAGAGCTGTGTGGGCTCTCCGAGCTGTTCCTCCCTGCACT 1225  
DB 246 CCACAGTGAAGATGAGTGGAGAGCTGTGTGGGCTCTCCGAGCTGTTCCTCCCTGCACT 1225  
QY 1226 CCAGCCCCCTTGAAGTGGAGATCTCAGGCTCTGTGAAAAACCTGGGAGAGTGGAGGCTCAG 1285  
DB 186 CCAGCCCCCTTGAAGTGGAGATCTCAGGCTCTGTGAAAAACCTGGGAGAGTGGAGGCTCAG 1285  
QY 1286 CTGCGGGGCAATGATTTGAGCCCTTGAAGAGAGATGAGGCTGCTTTGGAAGCCAGCA 1345

DB 126 CTGCGGGGCAATGATTTGAGCCCTTGAGGAGAGATAGGAGCTGCGCTTTGTGAAGCCAGCA 67  
QY 1346 GAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCTCCCACTTCTCTCTCTTA 1405  
DB 66 GAGGCTGAGAACCTCAGGCTTCCCTAGATCCAGAGCCCTCCCACTTCTCTCTCTTA 1405  
QY 1406 AAACAA 1411  
DB 6 AAAAAA 1

RESULT 2  
AL578980/c 914 bp mRNA linear EST 16-FEB-2001  
LOCUS AL578980 LTI\_NFL006.PL2 Homo sapiens cDNA clone GS0DK012YC23 3  
DEFINITION prime, mRNA sequence.  
ACCESSION AL578980  
VERSION AL578980.1 GI:12943576  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 914)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).  
FEATURES  
source location/Qualifiers  
1..914  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GS0DK012YC23"  
/clone\_id="LTI\_NFL006.PL2"  
/tissue\_type="Placenta"  
/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : [filang@lifetech.com](mailto:filang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 177 a 271 c 261 g 204 t 1 others  
ORIGIN  
Query Match 57.8%; Score 898.4; DB 9; Length 914;  
Best Local Similarity 99.6%; Pred. No. 1.2e-157;  
Matches 910; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
QY 490 GACACAGACCTGATGCTTCTTACCTGCGCGTACCGGAGGCGCCGACAGCCACTG 549  
DB 914 GACACAGACCTGATGCTTCTTACCTGCGCGTACCGGAGGCGCCGACAGCCACTG 549  
QY 550 GCAGCTCATGAAGAGTATTAACCTGCTGGAGAGCAGACAGTGCAGCGCTGCCAAGAG 609  
DB 854 GCAGCTCATGAAGAGTATTAACCTGCTGGAGAGCAGACAGTGCAGCGCTGCCAAGAG 609  
QY 610 GGA-CCAGTGTGAACCTTCTGTATGACAGAGAGCTGATTAAGACAGTGAAG 668  
DB 794 GGA-CCAGTGTGAACCTTCTGTATGACAGAGAGCTGATTAAGACAGTGAAG 668  
QY 669 ACATGGAATGAGTCTCGGAACATGAGCTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 728  
DB 734 ACATGGAATGAGTCTCGGAACATGAGCTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 728  
QY 729 AGATTGCGAGCTGGAACAGAACTGATTAAGTGCAGGCTGCTAGTGAACAGATCAG 788

```
Dh 674 AGATTCCGACCTGGACAGGAACTGCAATGAGTGGACGTGCTAGTGACACATCAGAC 615
Qy 789 GCATGAGCTCTCCGGAATTCGACCAACAGACACTGGAGTGGCGGGCCGATGATGTC 848
Db 614 GCATGAGCTCTCCGGAATTCGACCAACAGACACTGGAGTGGCGGGCCGATGATGTC 555
Qy 849 GGTA CTTGATGCGCTCGGCTGATGATGACCTGGGAGAGCAACCAAGATTAACGATTTG 908
Db 554 GGTA CTTGATGCGCTCGGCTGATGATGACCTGGGAGAGCAACCAAGATTAACGATTTG 495
Qy 909 ATGTGACCTGTCTTGGAGGGTCCGGCTGGAAGATATCCCGGAAACAGAGTGTATCA 968
Db 494 ATGTGACCTGTCTTGGAGGGTCCGGCTGGAAGATATCCCGGAAACAGAGTGTATCA 435
Qy 969 AGCTGAGAACAAAGGATATTTCTTCAATGCAATGAGAGTGCAGCGCCATCTACATCG 1028
Db 434 AGCTGAGAACAAAGGATATTTCTTCAATGCAATGAGAGTGCAGCGCCATCTACATCG 375
Qy 1029 ATGACCGCGCGTGTCTGTTGAGCTTCAATGAGCGCTCAGCAACAACTGTGTGAGAGA 1088
Db 374 ATGACCGCGCGTGTCTGTTGAGCTTCAATGAGCGCTCAGCAACAACTGTGTGAGAGA 315
Qy 1089 TCGCCAGCTTGGATTCGTCTTCTTATCAACAGAGCTCATGCTCATCAGGCTG 1148
Db 314 TCGCCAGCTTGGATTCGTCTTCTTATCAACAGAGCTCATGCTCATCAGGCTG 255
Qy 1149 AGGCTGCCAATCAACACAGAGTGAAGATGTTGGAGAGTGGAGAGTGGAGGCTCTCGGC 1208
Db 254 AGGCTGCCAATCAACACAGAGTGAAGATGTTGGAGAGTGGAGAGTGGAGGCTCTCGGC 195
Qy 1209 CTGTTTCCCTCCGACATCCAGCCCTTGAAGTGGAGACTGAGCTCTGAGAAAACTG 1268
Db 194 CTGTTTCCCTCCGACATCCAGCCCTTGAAGTGGAGACTGAGCTCTGAGAAAACTG 135
Qy 1269 GGCAGTGGAGAGCTCACTGCGGGCCATTTGATTTGAGCTTTGAGGAGATAGGGCTGG 1328
Db 134 GGCAGTGGAGAGCTCACTGCGGGCCATTTGATTTGAGCTTTGAGGAGATAGGGCTGG 75
Qy 1329 CTTTGTGAACCCGACAGAGCTGAGAACTCAGGCTTCCCTGATTCAGAGCCCTCC 1388
Db 74 CTTTGTGAACCCGACAGAGCTGAGAACTCAGGCTTCCCTGATTCAGAGCCCTCC 15
Qy 1389 CATCTTCTCTCTC 1402
Db 14 MATCTTCTCTCTC 1
```

```
RESULT 3
AL523921 927 bp mRNA linear EST 13-PEB-2001
LOCUS AL523921 LTI_NFL003.NBC3 Homo sapiens cDNA clone CS0DC003YA08 5
DEFINITION AL523921 LTI_NFL003.NBC3 Homo sapiens cDNA clone CS0DC003YA08 5
ACCESSION AL523921
VERSION AL523921.1 GI:12787414
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 927)
AUTHORS Li, W. B., Gruber, C., Jesssee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
```

## FEATURES

```
source
1..927
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC003YA08"
/clone_lib="LTI_NFL003_NBC3"
```

```
/sex="male"
/issue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact : Feng Liang life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 216 a 261 c 268 g 169 t 13 others
ORIGIN
Query Match 57.3%; Score 890; DB 9; Length 927;
Best Local Similarity 97.1%; Pred. No. 4.5e-156;
Matches 901; Conservative 13; Mismatches 13; Indels 1; Gaps 1;
Qy 163 TCAGGTGACCAAGATCTGGGCGCTGGAAGCTGCAATGACCTGCTCATTAATGC 222
Db 1 TCAGGTGACCAAGATCTGGGCGCTGGAAGCTGCAATGACCTGCTCATTAATGC 60
Qy 223 TGTGTGACAGCAACGACTGACCTCCGTCACTGGCGGTGAATTCAGCTGCCCTT 282
Db 61 TGTGTGACAGCAACGACTGACCTCCGTCACTGGCGGTGAATTCAGCTGCCCTT 120
Qy 283 CACCTTTGGGAGGTCCAGAGCGTTGATCGCTGTCTTACATCTGTATCTCCAA 342
Db 121 CACCTTTGGGAGGTCCAGAGCGTTGATCGCTGTCTTACATCTGTATCTCCAA 180
Qy 343 GTTGGGCTGTGACCAATGAGAGCTGCAACCAAGGCTATTGACAGCATCCAGAGCA 402
Db 181 GTTGGGCTGTGAGG-CATGAGAGAGCTGCAACCAAGGCTATTGAGCATCCAGAGCA 239
Qy 403 GGCCCTGTTTGAAGGCTGAGAGAGCTGCTGAGCAAAATGGAGATGCAACGAGCAAC 462
Db 240 GGCCCTGTTTGAAGGCTGAGAGAGCTGCTGAGCAAAATGGAGATGCAACGAGCAAC 299
Qy 463 CACCTTGAAGACTTTCAGAGCTGCTGACAGACACCTGATGCTTCTTACTGGCCG 522
Db 300 CACCTTGAAGACTTTCAGAGCTGCTGACAGACACCTGATGCTTCTTACTGGCCG 359
Qy 523 TACCGGCAAGGCCCTGACAGGCCCACTGGACACTATGAAGAGATTAACCTGCGAGAGA 582
Db 360 TACCGGCAAGGCCCTGACAGGCCCACTGGACACTATGAAGAGATTAACCTGCGAGAGA 419
Qy 583 CCAGACAGTGCAGCGCTGCCCAAGGGGACCAAGTGTGAACTTCTGTATGACAGAGA 642
Db 420 CCAGACAGTGCAGCGCTGCCCAAGGGGACCAAGTGTGAACTTCTGTATGACAGAGA 479
Qy 643 CCTGATTTGATGACAGTAACTCAAAGACATGCGAGATGAGTCTCTGAAACATGAGCTGAT 702
Db 480 CCTGATTTGATGACAGTAACTCAAAGACATGCGAGATGAGTCTCTGAAACATGAGCTGAT 539
Qy 703 GGTGGCTGACCGGGGCCCAAGGAGGATTCGGCACTGGAACAGAGAACTGCAATAGTG 762
Db 540 GGTGGCTGACCGGGGCCCAAGGAGGATTCGGCACTGGAACAGAGAACTGCAATAGTG 599
Qy 763 GCAGGTGCTAATGAGACAGATCAACAGGATGAGCTCTCCGAGCTTCGACCAACAGACACT 822
Db 600 GCAGGTGCTAATGAGACAGATCAACAGGATGAGCTCTCCGAGCTTCGACCAACAGACACT 659
Qy 823 GGCAGTGTGCGGGGCGCATGATGTCGGTACCTGATGAGCTCGCGTGAATCACCTTGGG 882
Db 660 GGCAGTGTGCGGGGCGCATGATGTCGGTACCTGATGAGCTCGCGTGAATCACCTTGGG 719
Qy 883 CAGAGCAACCAAGATTAACGATTTGATGATGAGCTGTCTTGAAGAGGTCCGGCTGGAA 942
Db 720 CAGAGCAACCAAGATTAACGATTTGATGATGAGCTGTCTTGAAGAGGTCCGGCTGGAA 779
Qy 943 GATATCCCGAAACAAAGGTGTCAATCAAGCTGAAGAACAAACGATGATTTCTTCAATTGCCAA 1002
```

Db 780 GATATCCGGAAACAAGTGTCTATCACTGAAGAAACAAGGATTTCTTCAATGCCAA 839  
Qy 1003 TGAGGATCGACGGCCCATCTACATCGATGACGGCCGGCTCTCTGTGCTCCAAATGCGC 1062  
Db 840 TGAGGATCGACGGCCCATCTACATCGATGACGGCCGGCTCTGTGCTCCAAATGCGC 899  
Qy 1063 CCTGACACAACTCTGTGTGAGATC 1090  
Db 900 CCTCAACACMACCTCTGTGTGAGATC 927

RESULT 4  
AL575639/c 993 bp mRNA linear EST 16-FEB-2001  
LOCUS AL575639 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1069YL09 3  
DEFINITION Prime, mRNA sequence.  
ACCESSION AL575639  
VERSION AL575639.1 GI:12936998  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 993)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers

FEATURES  
source 1..993  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1069YL09"  
/clone\_11b="LTI\_NFL006.PL2"  
/tissue\_type="Placenta"  
/note="Vector: PCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@life.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 206 a 276 c 285 g 220 t 6 others  
ORIGIN

Query Match 56.9%; Score 883.6; DB 9; Length 993;  
Best Local Similarity 96.1%; Pred. No. 6.6e-155;  
Matches 971; Conservative 4; Mismatches 14; Indels 21; Gaps 6;

Qy 488 CTGCACAGACACCTGATGCTCTTCTTCTGACCGGACCGGAGGCGCTGCAGGCCAC 547  
Db 990 CTGTGCAACACCTGATGCTCTTCTTCTGACCGGAGGCGGAGGCGCTGCAGGCCAC 933  
Qy 548 TGGCAGCTCATGAAGAGATTAATCTGCTGAGAGACAGACAGAGTGAAGCGGCGCCCAA 607  
Db 932 TGGCAGCTCATGAAGAGATTAATCTGCTGAGAGACAGACAGAGTGAAGCGGCGCCCAA 878  
Qy 608 GGGACCAAGTGTGAATCTCTGTATGACGAGGAGCTGATGACAGTAAGCTCAAG 667  
Db 877 -GGACCAAGTGTGAATCTCTGTATGACGAGGAGCTGATGACAGTAAGCTCAAG 819  
Qy 668 GACATGCGAGATGAGGTCTCTGAAACATGAGCTGATGCTGACCGGCGCCGAGAGCA 727  
Db 818 GACATGCGAGATGAGGTCTCTGAAACATGAGCTGATGCTGACCGGCGCCGAGAGCA 759  
Qy 728 GAGATTGCGAGCTGAGAACAGAACTGATAGTGGCAGGTGCTAGTGAACAGATCACA 787

Db 758 GATATCCGGACGTGGAACAGAACTGATTAAGTGGCAGGTCTAGTGGACAGATCA 699  
Qy 788 GGCATGAGCTCTCCGCACTTGGACAAACAGACATGGCAATGCTCGGGCGCGATGGT 847  
Db 698 GGCATGAGCTCTCCGCACTTGGACAAACAGACATGGCAATGCTCGGGCGCGATGGT 639  
Qy 848 CGGTACTGATGCGCTCGGTGAGATCAACCTGGGACAGCAACAGATTAACAGATT 907  
Db 638 CGGTACTGATGCGCTCGGTGAGATCAACCTGGGACAGCAACAGATTAACAGATT 579  
Qy 908 GATGAGACCTGTCTGTGAGGGTCCGGCTGGAAGATATCCCGAAACAAGTGTCAATC 967  
Db 578 GATGAGACCTGTCTGTGAGGGTCCGGCTGGAAGATATCCCGAAACAAGTGTCAATC 519  
Qy 968 AAGCTGAAGACACGAGTATTTCTTCAATGCAATGAGGGTGCAGCGCCATCTTCAATC 1027  
Db 518 AAGCTGAAGACACGAGTATTTCTTCAATGCAATGAGGGTGCAGCGCCATCTTCAATC 459  
Qy 1028 GATGAGACGGCGGTCTGTGAGCTCCAAATGGGCGCTGAGCAACATCTGTGGTGAAG 1087  
Db 458 GATGAGACGGCGGTCTGTGAGCTCCAAATGGGCGCTGAGCAACATCTGTGGTGAAG 399  
Qy 1088 ATGCGACGCTCGATGCTGTCTTCTTATCAACAGACCTTATGCTTCAATGAGGCT 1147  
Db 398 ATGCGACGCTCGATGCTGTCTTCTTATCAACAGACCTTATGCTTCAATGAGGCT 339  
Qy 1148 GAGGCTGCCAAGATCAACACAGTGAAGATGTGGAGACTGTGGGCGCTTCCGG 1207  
Db 338 GAGGCTGCCAAGATCAACACAGTGAAGATGTGGAGACTGTGGGCGCTTCCGG 279  
Qy 1208 CTTGTTTCCCGGCACTCCAGCCCTTGAAGCTGAGGAACTAGGCTCTGGAAGAACT 1267  
Db 278 CTTGTTTCCCGGCACTCCAGCCCTTGAAGCTGAGGAACTAGGCTCTGGAAGAACT 219  
Qy 1268 GGGCAGTGGAGGCTCAAGTGGGCGCATTTGAGCTTTGAGGAGATAGGAGCTG 1327  
Db 218 GGGCAGTGGAGGCTCAAGTGGGCGCATTTGAGCTTTGAGGAGATAGGAGCTG 159  
Qy 1328 GCTTTTGAACGACAGAGGCTGAGAACTCAAGCTTCTTATGATCCAGGCGCTCC 1387  
Db 158 GCTTTTGAACGACAGAGGCTGAGAACTCAAGCTTCTTATGATCCAGGCGCTCC 99  
Qy 1388 CCATTTCTCTCTCTTAAACAAACCTTACCCCATCTTACCCCATTTGACACTTCA 1447  
Db 98 CCGTCTCTCTCTCTTAAACAAACCTTACCCCATCTTACCCCATTTGACACTTCA 51  
Qy 1448 CTCCTGTGCTCAGCTGA-TTGGCTCAGACTCTTCTTTATGTTTTT 1496  
Db 50 CTCCTGTGCTCAGCTGANTTAGCTCAGACTCTTCTTTATGTTTTT 1

RESULT 5  
AL576040/c 913 bp mRNA linear EST 16-FEB-2001  
LOCUS AL576040 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1068YA02 3  
DEFINITION Prime, mRNA sequence.  
ACCESSION AL576040  
VERSION AL576040.1 GI:12937792  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 913)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers

FEATURES



```

source
1. 913
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="CS0D1068YA02"
/clone_lib="LTI_NFL006_PL12"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      190 a      257 c      260 g      205 t      1 others
ORIGIN

Query Match      56.4%; Score 876.4; DB 9; Length 913;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 910; Conservative 1; Mismatches 2; Indels 13; Gaps 2;

QY 566 TATTACCTGCTGAGACCAAGCAAGTGCAGCCGCTGCCAAAGGAGCAACCAAGTGTGAAC 625
DB 913 TATTACCTGCTGAGACCAAGCAAGTGCAGCCGCTGCCAAAGGAGCAACCAAGTGTGAAC 855
QY 626 TTCTCTGATGACAGAGACCTGATTGATGACATTAAGCTCAAGACATGCGAGATGAGTTC 685
DB 854 TTCTCTGATGACAGAGACCTGATTGATGACATTAAGCTCAAGACATGCGAGATGAGTTC 795
QY 686 CTGAACATGAGCTGATGATGCTGACCGCGCCCAAGACGAGAGATTGCGGACGTGAA 745
DB 794 CTGAACATGAGCTGATGATGCTGACCGCGCCCAAGACGAGAGATTGCGGACGTGAA 735
QY 746 CAGGAATGCAATGAGTGGCAGGTGCTAGTGAACAGATCAGAGCATGAGCTTCCGAGC 805
DB 724 CAGGAATGCAATGAGTGGCAGGTGCTAGTGAACAGATCAGAGCATGAGCTTCCGAGC 675
QY 806 TTGACCAACCAACACACTGGCACTGCTGCGGGCCGATGATGCTGATGCTGCTG 865
DB 674 TTGACCAACCAACACACTGGCACTGCTGCGGGCCGATGATGCTGATGCTGCTG 615
QY 866 CGTGAGATCAACCTGGGCGAGCAACCAAGATTAACCAATGATGTGACCTGTCTG 925
DB 614 CGTGAGATCAACCTGGGCGAGCAACCAAGATTAACCAATGATGTGACCTGTCTG 555
QY 926 GAGGATCCGCGCTGGAAGATATCCGGAACAAGGTGTCATCAAGCTGAAGAACACGCT 985
DB 554 GAGGATCCGCGCTGGAAGATATCCGGAACAAGGTGTCATCAAGCTGAAGAACACGCT 495
QY 986 GATTTCCTCATTTGCCAATGAGGGTGCACGGCCCATCTACATGATGAGACGCGCGTCTC 1045
DB 494 GATTTCCTCATTTGCCAATGAGGGTGCACGGCCCATCTACATGATGAGACGCGCGTCTC 435
QY 1046 TGTGCTCCAAATGGCGCTCGACCAACACTGCTGATGAGATGGCCGCTCGGATTC 1105
DB 434 TGTGCTCCAAATGGCGCTCGACCAACACTGCTGATGAGATGGCCGCTCGGATTC 375
QY 1106 GTCTTCCTTATCAACCCAGAGCTCATTTGCCCTATCATGAGGCTGAGGCTTCCCAAGATCA 1165
DB 374 GTCTTCCTTATCAACCCAGAGCTCATTTGCCCTATCATGAGGCTGAGGCTTCCCAAGATCA 315
QY 1166 CCACAGTAGAGATGTTGTCAGAGACTCTGTGGCCCTCTTCGCGCTGTTTCCCTGCGACT 1225
DB 314 CCACAGTAGAGATGTTGTCAGAGACTCTGTGGCCCTCTTCGCGCTGTTTCCCTGCGACT 255
QY 1226 CCAGCGCCCTTGAAGCTGGGAACTCAAGGCTCTGGAAGAACTGGGAGAGGAGGCTGAG 1285
DB 254 CCAGCGCCCTTGAAGCTGGGAACTCAAGGCTCTGGAAGAACTGGGAGAGGAGGCTGAG 195
QY 1286 CTGCGGGCCATTGATTGAGCTTTGAGGAGAGATGAGGCTGCTGTTGTGAAGCAGAG 1345

```

```

DB 194 CTGCGGGCCATTGATTGAGCTTTGAGGAGAGATAGGGCTGGCCCTTTGTGAAGCCAGCA 135
QY 1346 GAGGCTGAGAACCTTCAGGCTTCCCTAGAACAGAGCCCTCCCATCTTCTCTCTAA 1405
DB 134 GAGGCTGAGAACCTTCAGGCTTCCCTAGAACAGAGCCCTCCCATCTTCTCTCTAA 75
QY 1406 AAACAACCTTACCCCATCTTACCCCATCTTACCCCATCTTACCCCATCTTACCCCATCT 1465
DB 74 AAACAACCTTACCCCATCTTACCCCATCTTACCCCATCTTACCCCATCTTACCCCATCT 27
QY 1466 ATTACCTTCAAGACTTCTTTATTG 1491
DB 26 ATTACCTTCAAGACTTCTTTATTG 1

RESULT 6
AL581415/c
LOCUS      AL581415      1000 bp      mRNA      linear      EST 16-FEB-2001
DEFINITION AL581415 LTI_FLO11_BCl Homo sapiens cDNA clone CS0DG003Y101 3 prime
, mRNA sequence.
ACCESSION  AL581415
VERSION    AL581415
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1000)
            Li, W.B., Grubbs, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
COMMENT    Contact: Genoscope
            BP 191 91006 Evry cedex - France
            Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
            source
            1..1000
            /organism="Homo sapiens"
            /db xref="taxon:9606"
            /clone="CS0DG003Y101"
            /clone_lib="LTI_FLO11_BCl"
            /sex="male"
            /tissue_type="B cells from Burkitt lymphoma"
            /lab host="DH10B"
            /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched,
            double-stranded cDNA was digested with Not I and cloned
            into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
            Library was constructed by Life Technologies. Contact :
            Feng Liang Life Technologies, a division of Invitrogen
            9800 Medical Center Drive Rockville, Maryland 20850, USA
            Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com"

BASE COUNT      218 a      278 c      282 g      214 t      8 others
ORIGIN

Query Match      56.1%; Score 872; DB 9; Length 1000;
Best Local Similarity 94.1%; Pred. No. 9.5e-153;
Matches 947; Conservative 6; Mismatches 36; Indels 17; Gaps 4;

QY 500 CCTGATGCTTCTTACCTGCGCGTACCGGAAGGCTTCGAGGGCCCATGCGAGCTCATG 559
DB 990 CCTTATGCTTCTTA-CTTGCCCTTACCGGAAGGCC--TGACGCCCATGCGAGCTCATG 934
QY 560 AAGCAGATTACCTGCTGAGAGACCAAGCAAGTGCAGCCGCTGCCAAAGGAGCAACAGTG 619
DB 993 AAGCAGATTACCTGCTGAGAGACCAAGCAAGTGCAGCCGCTGCCAAAGGAGCAACAGTG 874
QY 620 CTGAACCTTCTGATGACAGAGACCTGATTGATGACATTAAGCTCAAGACATGAGAGAT 679
DB 873 CTGAACCTTCTGATGACAGAGACCTGATTGATGACATTAAGCTCAAGACATGAGAGAT 814
QY 680 GAGTCTCGAAACATGAGCTGATGCTGACCGCGCCCAAGACGAGAGATTGCGCAG 739

```



Qy 1070 AACAACTCTGTGTGAGATGCGCCAGCTGCGATTGCTTCTTATCAACGAGACCTC 1129  
Db 369 AACAACTCTGTGTGAGATGCGCCAGCTGCGATTGCTTCTTATCAACGAGACCTC 310  
Qy 1130 ATTGCCCTCATCAGGCGTGAAGCTGCCAAGATCACACCACTGAGTAATGTGTGCAGAGA 1189  
Db 309 ATTGCCCTCATCAGGCGTGAAGCTGCCAAGATCACACCACTGAGTAATGTGTGCAGAGA 250  
Qy 1190 CTCTGAGGACCTCTCTCGGACCTGTTTCCCTGACACTCCAGCCCCCTTGAAGCTGAGAACTC 1249  
Db 249 CTCTGAGGACCTCTCTCGGACCTGTTTCCCTGACACTCCAGCCCCCTTGAAGCTGAGAACTC 190  
Qy 1250 AGGCTCTGAAAAAAGCTGGGAGTGGAGAGCTCAGCTGCGGAGCTGATGATTTGAGCCTT 1309  
Db 189 AGGCTCTGAAAAAAGCTGGGAGTGGAGAGCTCAGCTGCGGAGCTGATGATTTGAGCCTT 130  
Qy 1310 TGAGGAGAGATGAGGCTGGCTTTTGTGAAGCCAGCAGAGGCTGAGAACTCAGGCTTCCC 1369  
Db 129 TGAGGAGAGATGAGGCTGGCTTTTGTGAAGCCAGCAGAGGCTGAGAACTCAGGCTTCCC 70  
Qy 1370 TAGATCCAGAGCCCTCCCATCTTCTCTCTTAAAAAACAACCTTACCCCATTT 1425  
Db 69 TAGATCCAGAGCCCTCCCATCTTCTCTCTTAAAAAACAACCTTACCCCATTT 14

RESULT 8  
AL544010/c 882 bp mRNA linear EST 16-FEB-2001  
LOCUS AL544010 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1004YM07 3  
DEFINITION AL544010 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1004YM07 3  
prime, mRNA sequence.  
ACCESSION AL544010  
VERSION AL544010.1 GI:12876489  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 882)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
Source  
1..882  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1004YM07"  
/clone\_lib="LTI\_NFL006.PL2"  
/tissue\_type="Placenta"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifestech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 175 a 251 c 253 g 199 t 4 others  
ORIGIN

Query Match 54.4%; Score 844.8; DB 9; Length 882;  
Best Local Similarity 97.6%; Pred. No. 1,2e-147;  
Matches 868; Conservative 3; Mismatches 7; Indels 11; Gaps 1;

Qy 578 GAGAGCCAGACAGTCAGCGCTGCCCAAGGAGCAAGTCTTAATCTTCTGATGCA 637  
Db 878 GAGAGCCAGACAGTCAGCGCTGCCCAAGGAGCAAGTCTTAATCTTCTGATGCA 819

Qy 638 GAGAGCTGATTTATGACGTAAGCTCAAGGACATGCGAGATGAGTCTTGAACATGAG 697  
Db 818 GAGAGCTGATTTATGACGTAAGCTCAAGGACATGCGAGATGAGTCTTGAACATGAG 759  
Qy 698 CTGATGTGGCTGACCGGCGGACAGAGAGATTTGGCGAGCTGGAAACAGAACTGAT 757  
Db 758 CTGATGTGGCTGACCGGCGGACAGAGAGATTTGGCGAGCTGGAAACAGAACTGAT 699  
Qy 758 AAGTGGCAGGTGCTAGTGAAGAGATCAAGGACATGAGCTCTCGGAACTTGAACAACAG 817  
Db 698 AAGTGGCAGGTGCTAGTGAAGAGATCAAGGACATGAGCTCTCGGAACTTGAACAACAG 639  
Qy 818 AACTGGAGTGTGCGGAGCGCGATGAGTGCGTAACCTGATGCGCTGCGTGAATCAC 877  
Db 638 AACTGGAGTGTGCGGAGCGCGATGAGTGCGTAACCTGATGCGCTGCGTGAATCAC 579  
Qy 878 CTGGGCAAGCAACCAAGATTAACGATTTGATGTGAACCTGTCTTGGAGAGTCCGGCC 937  
Db 578 CTGGGCAAGCAACCAAGATTAACGATTTGATGTGAACCTGTCTTGGAGAGTCCGGCC 519  
Qy 938 TGGAAGATATCCCGGAAACAGAGTGTCAATCAAGCTGAAGAACAAACGATTTCTTCAAT 997  
Db 518 TGGAAGATATCCCGGAAACAGAGTGTCAATCAAGCTGAAGAACAAACGATTTCTTCAAT 459  
Qy 998 GCCAATGAGGCTGACAGCGCCCATCTACATGATGAGACGCGGATCTCTGTGCTCCAAA 1057  
Db 458 GCCAATGAGGCTGACAGCGCCCATCTACATGATGAGACGCGGATCTCTGTGCTCCAAA 399  
Qy 1058 TGGCGCTCAGCAACAACTCTGTGTGAGATCGCCAGCTGCGATTTGCTTCTTATC 1117  
Db 398 TGGCGCTCAGCAACAACTCTGTGTGAGATCGCCAGCTGCGATTTGCTTCTTATC 339  
Qy 1118 AACCAAGACCTCATTCGCCCTCATGAGGCTGAGGCTGCAAGATCACACCAAGAGAGA 1177  
Db 338 AACCAAGACCTCATTCGCCCTCATGAGGCTGAGGCTGCAAGATCACACCAAGAGAGA 279  
Qy 1178 ATGTGGCAGAGCTCTGTGGGCTCTCTCGGCTGTTTCCCTGCACTCCAGCCCTTGG 1237  
Db 278 GTGTGGCAGAGCTCTGTGGGCTCTCTCGGCTGTTTCCCTGCACTCCAGCCCTTGG 219  
Qy 1238 AGCTGGGAACCTCAGGCTCTGGAAGAACCTGGGAGCTGAGGCTCAGCTCGGGCCATT 1297  
Db 218 AGCTGGGAACCTCAGGCTCTGGAAGAACCTGGGAGCTGAGGCTCAGCTCGGGCCATT 159  
Qy 1298 GATTGAGCCTTTGAGGAGATAGAGGCTGTCCTTTGGAAGCCAGCAGAGCTGAGAAC 1357  
Db 158 GATTGAGCCTTTGAGGAGATAGAGGCTGTCCTTTGGAAGCCAGCAGAGCTGAGAAC 99  
Qy 1358 CTGAGGCTTCCCTAGATCAGAGCCCTCCCATCTTCTCTTAAAAAACAACCTTAC 1417  
Db 98 CTGAGGCTTCCCTAGATCAGAGCCCTCCCATCTTCTCTTAAAAAACAACCTTAC 42  
Qy 1418 CCCCCATTCTAACCCCATTTGCCAATCTCCTGTGTCTCCAGCTGA 1466  
Db 41 CCCCCATTCTAACCCCATTTGCCAATCTCCTGTGTCTCCAGCTGA 1

RESULT 9  
BMS49097 1021 bp mRNA linear EST 20-FEB-2002  
LOCUS BMS49097  
DEFINITION AGENCOURT 6554149 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5548420  
5', mRNA sequence.  
ACCESSION BMS49097  
VERSION BMS49097.1 GI:18784222  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1021)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsb@rs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: LLM1257 row: F column: 05  
High quality sequence stop: 731.  
Location/Qualifiers

FEATURES  
source  
1..1021  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5548420"  
/clone\_lib="NIR-MGC\_72"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 232 a 292 c 292 g 195 t 10 others

ORIGIN

Query Match 53.9%; Score 836.6; DB 13; Length 1021;  
Best Local Similarity 96.8%; Pred. No. 3.7e-146;  
Matches 907; Conservative 0; Mismatches 21; Indels 9; Gaps 5;

151 TAAACGACCACTTCAAGTACCAAGATCTGGAGCCGCTGGAAGCTGCAATGACCTCT 210  
Db 1 TAAACGACCACTTCAAGTACCAAGATCTGGAGCCGCTGGAAGCTGCAATGACCTCT 60

211 GCTCAATAATGCTGTGTTCAGACCAACGACCTGACCTCGTCCACCTGGGCGTGAATT 270  
Db 61 GCTCAATAATGCTGTGTTCAGACCAACGACCTGACCTCGTCCACCTGGGCGTGAATT 120

271 CAGCTGCCGCTTCACTTCGGGAGGTCCAGAGAGCTGTGTACGCTCTCTTACGATCC 330  
Db 121 CAGCTGCCGCTTCACTTCGGGAGGTCCAGAGAGCTGTGTACGCTCTCTTACGATCC 180

331 TGTCAATCTCAAGTGGCGCTGACAGGCGATGAGGCGAGCTGACCCGAGAGCTATTGACG 390  
Db 181 TGTCAATCTCAAGTGGCGCTGACAGGCGATGAGGCGAGCTGACCCGAGAGCTATTGACG 240

391 CATTCAAGACCAAGGCGCTTTAGCAAGGCTGAGAGAGCTGCTAGCAAAATGGGATC 450  
Db 241 CATTCAAGACCAAGGCGCTTTAGCAAGGCTGAGAGAGCTGCTAGCAAAATGGGATC 300

451 GACCAAGCCGCACTTGTGAGACCTTCCAGACCTGCTGACAGACACCTTGATGCTTT 510  
Db 301 GACCAAGCCGCACTTGTGAGACCTTCCAGACCTGCTGACAGACACCTTGATGCTTT 360

511 CTAACGAGCCGCTGACCGGAGGCGCTGACAGGCGCACTGAGAGCTATGAAGAGATT 570  
Db 361 CTAACGAGCCGCTGACCGGAGGCGCTGACAGGCGCACTGAGAGCTATGAAGAGATT 420

571 CCGTGTGAGGACCAAGACAGTGCAGCGCTGCCCAAGGGGACCAAGTCTGAATCTTC 630  
Db 421 CCGTGTGAGGACCAAGACAGTGCAGCGCTGCCCAAGGGGACCAAGTCTGAATCTTC 480

631 TGAATGACAGGACCTGATTTGATGACAGTAAAGCTCAAGACATGCGAGATGAGTCTTGA 690  
Db 481 TGAATGACAGGACCTGATTTGATGACAGTAAAGCTCAAGACATGCGAGATGAGTCTTGA 540

691 AATAGAGTGAATGAGTGGCGGACCGGCGCAAGAGAGAGATTCGCGACCTGGAACAGA 750  
Db 541 AATAGAGTGAATGAGTGGCGGACCGGCGCAAGAGAGAGATTCGCGACCTGGAACAGA 600

751 ACTGCAATGAGGAGTCTAGTGAAGACATCAGAGCATAGAGTCTCCGACTTCCA 810  
Db 601 ACTGCAATGAGGAGTCTAGTGAAGACATCAGAGCATAGAGTCTCCGACTTCCA 660

811 CAACGACACCTGAGGAGTGTGCGGGCGGACATGAGTGGGTAATGATGCGTCCGCTGA 870  
Db 661 CAACGACACCTGAGGAGTGTGCGGGCGGACATGAGTGGGTAATGATGCGTCCGCTGA 720

871 GATCAACCTGGGACAGGACCAACCAAGATTAACCAATGATGAGCTGTCTTGGAGGG 930  
Db 721 GATCAACCTGGGACAGGACCAACCAAGATTAACCAATGATGAGCTGTCTTGGAGGG 780

931 TCCGGCTTGGAGAGATATCCG-GGAAACAGGTGTATCAAGTGG-AAAGAACAGGTAT 988  
Db 781 TCCGGCTTGGAGAGATATCCG-GGAAACAGGTGTATCAAGTGGAAAGCAACAGGTAT 840

989 TTCTTCATTTGCCATGA-GGGTCAGGCGCCCATCTACA-TCGATGAGAGGCC-----GGT 1041  
Db 841 TTCTTCATTTGCCATGAAGGATGAGGTCAGCGCCCATCTACANTGATGAGAGCGCGGTGCT 900

1042 GCTCTGTGGCTCCAAATGAGCGGCTGAGCAACAACCTT 1078  
Db 901 CCTGTGGCTCCCAATGAGGCGGCTGAGCAACAACCT 937

RESULT 10  
AL544052 848 bp mRNA linear EST 16-FEB-2001  
LOCUS AL544052 LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1004YM07 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL544052  
VERSION AL544052.1 GI:12876531  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 848)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source  
1..848  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1004YM07"  
/clone\_lib="LTI.NFL006.PL2"  
/tissue\_type="Placenta"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime and enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 202 a 235 c 252 g 157 t 2 others

ORIGIN

Query Match 53.6%; Score 833; DB 9; Length 848;  
Best Local Similarity 99.4%; Pred. No. 1.9e-145;  
Matches 844; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

333 TCATCTCCAAAGTTGGCTGTGACGACATGAGGACCTCAACCAAGAGCTATTGAGCA 392  
Db 1 TCATCTCCAAAGTTGGCTGTGACGACATGAGGACCTCAACCAAGAGCTATTGAGCA 60

393 TCAGAGCAAGGCGCTGTTTGAAGGCTGAGAGCACTGCTGAGCAAAATGGGATGCA 452  
Db 61 TCAGAGCAAGGCGCTGTTTGAAGGCTGAGAGCACTGCTGAGCAAAATGGGATGCA 120

QY 453 CCAGCAGCCACCTTGGAGACTTCCAGAGCTGCTCAGAGACACCTGATGCTTCT 512  
 DB 121 CCAAGCAGCCACCTTGGAGACTTCCAGAGCTGCTCAGAGACACCTGATGCTTCT 180  
 QY 513 ACCTGACCCGCTACCGGAGAGCCCTGACAGGCCACTGCGAGCTCATGAGACGATTTACC 572  
 DB 181 ACCTGACCCGCTACCGGAGAGCCCTGACAGGCCACTGCGAGCTCATGAGACGATTTACC 240  
 QY 573 TGCAGAGAGACCAAGCACTGACCGGCTGCCCAAGGGGACCAAGTGTGAATCTTCTG 632  
 DB 241 TGCAGAGAGACCAAGCACTGACCGGCTGCCCAAGGGGACCAAGTGTGAATCTTCTG 300  
 QY 633 ATGACAGAGACCTGATTTGATGACAGTAACTCAAGAGATGCGAGATGAGGCTCTGAAAC 692  
 DB 301 ATGACAGAGACCTGATTTGATGACAGTAACTCAAGAGATGCGAGATGAGGCTCTGAAAC 360  
 QY 693 ATGAGCTGATGCTGCTGACCGGCGCCAGAAAGCGAGAGATTTGCGAGCTGGAACAGAAC 752  
 DB 361 ATGAGCTGATGCTGCTGACCGGCGCCAGAAAGCGAGAGATTTGCGAGCTGGAACAGAAC 420  
 QY 753 TGCATTAATGCGAGGCTGCTAGTGGACAGCATCAAGCATGAGTCTCTCCGACTTCAGACA 812  
 DB 421 TGCATTAATGCGAGGCTGCTAGTGGACAGCATCAAGCATGAGTCTCTCCGACTTCAGACA 480  
 QY 813 ACCAGACACTGCGAGTGTGCGGGGCGGATGCTGCGGACTGATGCGCTGCGCTGAGA 872  
 DB 481 ACCAGACACTGCGAGTGTGCGGGGCGGATGCTGCGGACTGATGCGCTGCGCTGAGA 540  
 QY 873 TCACCTGCGGAGAGCAACCAAGATTAACCAAGATTGATGAGACTGCTCTCTGAGAGGCTC 932  
 DB 541 TCACCTGCGGAGAGCAACCAAGATTAACCAAGATTGATGAGACTGCTCTCTGAGAGGCTC 600  
 QY 933 CGGCTGGAAGATATCCCGGAAACCAAGGTGTATCAAGCTGAAGAAACAACGGTGAATTTCT 992  
 DB 601 CGGCTGGAAGATATCCCGGAAACCAAGGTGTATCAAGCTGAAGAAACAACGGTGAATTTCT 660  
 QY 993 TCATTGCAATGAGGGGTGACGGGCGCATCTACATGATGAGAGCGCGGAGCTCTGAGCT 1052  
 DB 661 TCATTGCAATGAGGGGTGACGGGCGCATCTACATGATGAGAGCGCGGAGCTCTGAGCT 720  
 QY 1053 CCAATGAGCGCTCAGCAACAACCTGTGTGTGAGAGATGCGCAGCTCGGATTCGTCTTC 1112  
 DB 721 CCAATGAGCGCTCAGCAACAACCTGTGTGTGAGAGATGCGCAGCTCGGATTCGTCTTC 779  
 QY 1113 TTATCAACCAAGACCTCATTTGCTTCATCAGGGCTGAGGCTGCCAAGATCACACCAAGT 1172  
 DB 780 TTATCAACCAAGACCTCATTTGCTTCATCAGGGCTGAGGCTGCCAAGATCACACCAAGT 839  
 QY 1173 GAGGAATGG 1181  
 DB 840 GAGGAATGG 848

RESULT 11  
 AL527295/c 947 bp mRNA linear EST 13-FEB-2001  
 LOCUS AL527295 LTI\_NFL003\_NBC3 Homo sapiens cDNA clone CS0DC021YD22 3  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL527295  
 VERSION AL527295.1 GI:12790788  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 947)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France

FEATURES  
 source  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 Location/Qualifiers  
 1..947  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="GS0DC021YD22"  
 /clone\_lib="LTI\_NFL003\_NBC3"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Peng liang life technologies, a division of invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 186 a 267 c 273 g 214 t 7 others  
 ORIGIN

Query Match 53.3%; Score 827.4; DB 9; Length 947;  
 Best Local Similarity 96.8%; Pred. No. 2e-144;  
 Matches 884; Conservative 6; Mismatches 8; Indels 15; Gaps 4;

QY 561 AGCAGTATTAATCTGCTGGAGAGCCAGACAGTGCAGCCGCTGCCAAGAGGACCAAGTGC 620  
 DB 900 AGCAGTATTAATCTGCTGGA -ACACAGACAGTGCAGCCGCTGCCAAGAGTGC 843  
 QY 621 TGAATCTCTGATGACAGAGACCTGATTTGACAGTAACTCAAGAGATGCGAGATG 680  
 DB 842 TGAATCTCTGATGACAGAGACCTGATTTGACAGTAACTCAAGAGATGCGAGATG 783  
 QY 681 AGGCTCTGGAACATGAGCTGATGTGCTGACCGGCGCCAGAAAGCGAGATTTGCGAGC 740  
 DB 782 AGGCTCTGGAACATGAGCTGATGTGCTGACCGGCGCCAGAAAGCGAGATTTGCGAGC 723  
 QY 741 TGGAAACAGAACTGATTAAGTGGAGGCTGATGAGACAGATCAACAGGCTGAGCTTC 800  
 DB 722 TGGAAACAGAACTGATTAAGTGGAGGCTGATGAGACAGATCAACAGGCTGAGCTTC 663  
 QY 801 CGGACTTGAACAACAGACACTGGCAGTGGCGGGCGCATGCTGGATGACCTGATGC 860  
 DB 662 CGGACTTGAACAACAGACACTGGCAGTGGCGGGCGCATGCTGGATGACCTGATGC 603  
 QY 861 GCTGCGTGAAGTACCTCTGGGAGAGCAACCAAGATTAACCAAGTTGATGAGCTGT 920  
 DB 602 GCTGCGTGAAGTACCTCTGGGAGAGCAACCAAGATTAACCAAGTTGATGAGCTGT 543  
 QY 921 CTCTGAGAGGTCGGGCTCTGGAAGATATCCGGAAACAAGGTGTATCAAGCTGAAGAAC 980  
 DB 542 CTCTGAGAGGTCGGGCTCTGGAAGATATCCGGAAACAAGGTGTATCAAGCTGAAGAAC 483  
 QY 981 ACGGTGATTTCTTCAATGCGCAATGAGGCTGACGGCCCATATCATGATGAGAGCGCGG 1040  
 DB 482 ACGGTGATTTCTTCAATGCGCAATGAGGCTGACGGCCCATATCATGATGAGAGCGCGG 423  
 QY 1041 TGCTCTGTGCTCCAAATGAGGCTCAGCAACAACCTGTGTGTGAGATTCGACGCTGC 1100  
 DB 422 TGCTCTGTGCTCCAAATGAGGCTCAGCAACAACCTGTGTGTGAGATTCGACGCTGC 363  
 QY 1101 GATTGCTTTCTTATCAACAGAGACCTGATTCCTCATCAGGGCTGAGGCTGCAAGA 1160  
 DB 362 GATTGCTTTCTTATCAACAGAGACCTGATTCCTCATCAGGGCTGAGGCTGCAAGA 303  
 QY 1161 TCACACCAAGTGAAGATGATGAGAGGACTGAGGCGCTCTCGGCTGATTTCCCTGT 1220  
 DB 302 TCACACCAAGTGAAGATGATGAGAGGACTGAGGCGCTCTCGGCTGATTTCCCTGT 243  
 QY 1221 CCACTCCAGCCCTTGAAGTGGAACTCAGGCTCTTGGAAAAAAGCTGGAGTGGAGG 1280

Db 242 CCACCTCCAGCCCCCTTGAGCTGGGAACCTCAGGCTCTGGAAAAACCTGGGCACTGGAGS 183  
 Qy 1281 CTCAGCTGGGGCCATTGATTTGAGCTTTGAGGAGATAGGGCTGGCTTTGTGAAC 1340  
 Db 182 CTCAGCTGGGGCCATTGATTTGAGCTTTGAGGAGATAGGGCTGGCTTTGTGAAC 123  
 Qy 1341 CAGCAGAGCTGAGAACCTCAGAGCTTACCCTAG--ATCCAGAGAGCCCTCCCACTTCTC 1398  
 Db 122 CAGCAGAGCTGAGAACCTCAGAGCTTCCWAGCATSSGAGCCCTCCCACTTCTC 63  
 Qy 1399 TCTTAAAAACAACCTTACCCCACTTCTACCCCACTTCTCACTCTGTGTCT 1458  
 Db 62 TCTTAAAAACAACCTTAC-----CGCCCATTTGCCACATCACTCTGTGTCT 14  
 Qy 1459 CCAGCTGATTAGC 1471  
 Db 13 CCAGCTGATTAGC 1

RESULT 12  
 AL529425/c 852 bp mRNA linear EST 13-FEB-2001  
 LOCUS AL529425 LTI\_NFL001\_NBC4 Homo sapiens cDNA clone CSDD006YA05 3  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL529425 GI:12792918  
 VERSION AL529425.1 GI:12792918  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 852)  
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
 source  
 Location/Qualifiers  
 1..852  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CSDD006YA05"  
 /clone\_11b="LTI\_NFL001\_NBC4"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liliang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 163 a 262 c 230 g 180 t 17 others  
 ORIGIN

Query Match 53.0%; Score 823.4; DB 9; Length 852;  
 Best Local Similarity 96.6%; Pred. No. 1.2e-143;  
 Matches 823; Conservative 16; Mismatches 13; Indels 0; Gaps 0;

Qy 492 ACAGACACCTGATGCTTCTACCTGCGCGTACCGGAGAGCGCTGACGAGCCCACTGAC 551  
 Db 852 ACARACACCTGATGCTTCTACCTGCGCGTACCGGAGAGCGCTGACGAGCCCACTGAC 793  
 Qy 552 ACCTCATGAGAGATTAATCTGCTGAGAGACCGACAGTGCAGCGCTGCGCAAGAGG 611  
 Db 792 MCTCATGAGAGATTAATCTGCTGAGAGACCGACAGTGCAGCGCTGCGCAAGAGG 733

Qy 612 ACCAGTGTGAATCTTCTGTGATGAGAGACCTGATTGATGACATGAGCTACAGACA 671  
 Db 732 ACCAGTGTGAATCTTCTGTGATGAGAGACCTGATTGATGACATGAGCTACAGACA 673  
 Qy 672 TGGAGATGAGGCTCTGGAACATGAGCTGATGATGATGATGATGATGATGATGATGAT 731  
 Db 672 TGGAGATGAGGCTCTGGAACATGAGCTGATGATGATGATGATGATGATGATGATGAT 613  
 Qy 732 TGGAGATGAGGCTCTGGAACATGAGCTGATGATGATGATGATGATGATGATGATGAT 791  
 Db 612 TGGAGATGAGGCTCTGGAACATGAGCTGATGATGATGATGATGATGATGATGATGAT 553  
 Qy 792 TGAGCTTCCGAGCTTTCACCAACAGACATGAGCTGAGCTGAGGAGCGGATGAGCT 851  
 Db 552 TGAGCTTCCGAGCTTTCACCAACAGACATGAGCTGAGCTGAGGAGCGGATGAGCT 493  
 Qy 852 ACCTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 911  
 Db 492 RGTGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 433  
 Qy 912 TGAGCTTCTTGTGAGAGGCTCGGCTGAGAGATATCCCGGAAACAAGTGTATCATAGC 971  
 Db 432 TGAGCTTCTTGTGAGAGGCTCGGCTGAGAGATATCCCGGAAACAAGTGTATCATAGC 373  
 Qy 972 TGAAGAACAGAGGATTTCTTCAATGCAATGAGAGGCTGAGAGGCTGATCATGATG 1031  
 Db 372 TGAAGAACAGAGGATTTCTTCAATGCAATGAGAGGCTGAGAGGCTGATCATGATG 313  
 Qy 1032 GACGAGCGGCTGCTGTGAGCTTCAATGAGCGCTGAGAACACTCTGTGTGAGATG 1091  
 Db 312 GACGAGCGGCTGCTGTGAGCTTCAATGAGCGCTGAGAACACTCTGTGTGAGATG 253  
 Qy 1092 CAGGCTGAGATCTCTTCTTCAATGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 1151  
 Db 252 CAGGCTGAGATCTCTTCTTCAATGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 193  
 Qy 1152 CTGCAAGATGACACACAGTGAAGATGATGAGAGGCTGAGAGGCTGAGAGGCTG 1211  
 Db 192 CTGCAAGATGACACACAGTGAAGATGATGAGAGGCTGAGAGGCTGAGAGGCTG 133  
 Qy 1212 TTTCCCTGCACTCTCAGAGCCCTTGAAGCTGGAATCAGAGCTCTGAAAAAAGCTGGGC 1271  
 Db 132 TTTCCCTGCACTCTCAGAGCCCTTGAAGCTGGAATCAGAGCTCTGAAAAAAGCTGGGC 73  
 Qy 1272 AGTGAAGGCTGAGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 1331  
 Db 72 AGTGAAGGCTGAGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 13

RESULT 13  
 LOCUS AL571554/c 916 bp mRNA linear EST 16-FEB-2001  
 DEFINITION AL571554 LTI\_NFL006\_P12 Homo sapiens cDNA clone CSDD0124YD06 3  
 prime, mRNA sequence.  
 ACCESSION AL571554  
 VERSION AL571554.1 GI:12928964  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 916)  
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.





```
Db 775 ACCGGGCGCCAGAGAGGATTCGGCACTGGAAACAGAACTGCAATKAGGCAAGTGC 716
Oy 771 TAGTGAACGACATTCACAGGCAATGAGCTCTCCGAGCTTCGACAA-CCAGACACTGGCAGTG 829
Db 715 TMTGTGACAGCATCAACAGGCAATGAGCTCTCCGACTTCCAGCAATCCAGACACTGGCAGTG 656
Oy 830 CCGCGGGGGCGGATGGTGGTGGTACCTGATGGCTGGCGTGGAGATCAACCTGGGCGAGAGA 889
Db 655 CTGCGGGG-CGCAATGGTGGGATACATKATGCGTTCCCGTGAAGATCACTGGGTAAGAGA 597
Oy 890 ACCAAGATTAACAGATTGATGATGAGCTGTCTCTGGAAGGTCGGGCTGGA-AGATATC 948
Db 596 ACYAAAGATTAACAGATTGATGAGCTGTCTCTGGAAGGTCGGGCTGGAATGATATC 537
Oy 949 CCGGAAACAGAGTGTCTATCAAGCTGAAGAACACGCTGATTTCTTATTGCCAATAGGG 1008
Db 536 CCGGAAATTAAGTGTCTATCAAGCTGAAGAACACGCTGATTTCTTATTGCCAATAGGG 477
Oy 1009 TCGAGCGGCCATCTACATGATGAGAGCGCGGCTGTGGTCCAAATGGGCGCTCAG 1068
Db 476 TCGAGCGGCCATCTACATGATGAGAGCGCGGCTGTGGTCCAAATGGGCGCTCAG 417
Oy 1069 CAACAACCTGTGTGATGAGATGCGCAGCTGGGATTCGTTCTTATCAACAGGACCT 1128
Db 416 CAACAACCTGTGTGATGAGATGCGCAGCTGGGATTCGTTCTTATCAACAGGACCT 357
Oy 1129 CATTGCGCTCATCAGAGGCTGAGAGCTGCCAAGATCACACACAGTGAAGATGTGGCAG 1188
Db 356 CATTGCGCTCATCAGAGGCTGAGAGCTGCCAAGATCACACACAGTGAAGATGTGGYAG 297
Oy 1189 ACTCGTGGGCGCTCTCGCGGCTGTTCCTCCCTGCACTCCAGCCCGCTTGA-GCTGGGAG 1247
Db 296 ACTGTGGGCGCTCTCTCGCGGCTGTTCCTCCCTGCACTCCAGCCCGCTTGAAGTGGGAG 237
Oy 1248 TCAGGCTCTCTGAAAAAACCCTGGGCACTGGAGGCTCAG-CTGGGCGCAATTGATTGAG 1306
Db 236 TCAGGCTCTCTGAAAAAACCCTGGGCACTGGAGGCTCAGCTGGGCGCAATTGATTGAG 177
Oy 1307 CTTTGAAGGAGATGAGGCTGGCTTTTGTGAACCGACAGAGGCTGAAGAACTCAGGCTT 1366
Db 176 CTTTGAAGGAGATGAGGCTGGCTTTTGTGAACCGACAGAGGCTGAAGAACTCAGGCTT 117
Oy 1367 CCTGATGATCAGAGCCGCCCATCTTCTCTGTAATAAACCAACCCCAACCCCATTC 1426
Db 116 CCTGATGATCAGAGCCGCCCATCTTCTCTGTAATAAACCAAC-----C 68
Oy 1427 TACCCCGCATTCACCTTCACCTCTGTGTCTCAGAGTATTAGCCTCAGACTCTTCTT 1486
Db 67 TACCCCGCATTCACCTTCACCTCTGTGTCTCAGAGTATTAGCCTTAACTCTCTTAT 8
Oy 1487 TATTGT 1492
Db 7 TTTT 2

RESULT 15
AL562353 860 bp mRNA linear EST 16-FEB-2001
LOCUS AL562353 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC007YE11 3
DEFINITION prime, mRNA Sequence.
ACCESSION AL562353
VERSION AL562353.1 GI:12910688
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 860)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
```

```
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
    source
        1..860
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="CS0DC007YE11"
            /clone_lib="LTI_NFL003_NBC3"
            /sex="male"
            /tissue_type="neuroblastoma cells"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-Oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Contact : Feng Liang Life
Technologies, a division of invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 160 a 259 c 234 g 199 t 8 others
ORIGIN
Query Match 52.3%; Score 812.4; DB 9; Length 860;
Best Local Similarity 96.6%; Pred. No. 1.3e-141;
Matches 831; Conservative 8; Mismatches 20; Indels 1; Gaps 1;

Oy 489 TGCACAGACACCCCTAGTGCCTTCTTCTGAGCCGCGGAGGCGCTGAGGCCACT 548
Db 860 TGAACAGACACCCCTAGTGCCTTCTTCTGAGCCGCGGAGGCGCTGAGGCCACT 801
Oy 549 GGCAGCTCATGAAGATTAATCTGCTGAGAGACAGACAGTGAAGCGCGCCCAAG 608
Db 800 GGCAGCTCATGAAGATTAATCTGCTGAGAGACAGACAGTGAAGCGCGCCCAAG 741
Oy 609 GGGACCAAGTGTGAACCTCTCTGATGACAGAGACCTGATTGATGACAGTAACTCAAG 668
Db 740 GGGACCAAGTGTGAACCTCTCTGATGACAGAGACCTGATTGATGACAGTAACTCAAG 681
Oy 669 ACATCGAGATGAGTCTCTGGAACATGAGCTGATGAGTGGCTGACCGGCGCAAGCGAG 728
Db 680 ACATCGAGATGAGTCTCTGGAACATGAGCTGATGAGTGGCTGACCGGCGCAAGCGAG 621
Oy 729 AGATTGCGCAGCTGGAACAGAACTGCATAAGTGCAGGTGTATGACACAGATCAAG 788
Db 620 AGATTGCGCAGCTGGAACAGAACTGCATAAGTGCAGGTGTATGACACAGATCAAG 561
Oy 789 GCATGAGCTCTCGGACTTCCGACCAACAGACACTGSCAGTGTGGGGGCGCATGTGC 848
Db 560 GCATGAGCTCTCGGACTTCCGACCAACAGACACTGSCAGTGTGGGGGCGCATGTGC 501
Oy 849 GGTACTGATGCGCTCGGTGATCACTCGGCGAGCAACCAAGATTAACAGATTG 908
Db 500 GGTACTGATGCGCTCGGTGATCACTCGGCGAGCAACCAAGATTAACAGATTG 441
Oy 909 ATGTGAGCTGTCTCTGAGAGGTCGGGCTTGAAGATATCCCGGAAACAGAGTGTCAATCA 968
Db 440 ATGTGAGCTGTCTCTGAGAGGTCGGGCTTGAAGATATCCCGGAAACAGAGTGTCAATCA 381
Oy 969 AGCTGAAGAACAGGATATTCTTCATTGCAATGAGAGTGTGACGGGCCATCTCATCTG 1028
Db 380 AGCTGAAGAACAGGATATTCTTCATTGCAATGAGAGTGTGACGGGCCATCTCATCTG 321
Oy 1029 ATGACGCGCGGTGCTCTGTGCTCAATGAGCGGCTCAGCAACCAACTCTGTGTGAGAGA 1088
Db 320 ATGACGCGCGGTGCTCTGTGCTCAATGAGCGGCTCAGCAACCAACTCTGTGTGAGAGA 261
Oy 1089 TCGCAGGCTGAGATTGCTTCTTATCAACAGAGACTCATTTGCCCTCATCAGGCTG 1148
Db 260 TCGCAGGCTGAGATTGCTTCTTATCAACAGAGACTCATTTGCCCTCATCAGGCTG 201
```



```

Qy 1149 AGGCTGCCAAGATACACCCAGTGGAGAAATGGTGACAGACTCGTGGGCCCTCTCCGGC 1208
    |||||
Db 200 AGGCTGCCAAGATACACCCAGTGGAGAAATGGTGACAGACTCGTGGGCCCTCTCCGGC 141
    |||||
Qy 1209 CTGTTTCCCCCTGCCACTCCAGCCCCCTTGAGCTGGAACTCAGGCTCCTGGAAAAACCTG 1268
    |||::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 140 CTGAAAMWCTGCMACCTCAGCCCCCHTGAAGCTGGGACTCAGGCTCCTGGAAAAACCTG 81
    |||||
Qy 1269 GGCAGTGGAGGCTCAGCTGGGGCCATTGATTTGAGCCCTTGAGGGAGAGATAGG-GCTG 1327
    |||||
Db 80 GGCAGTGGAGGCTCAGCTGGGGCCATTGATTTGAGCCCTTGAGGGAGAGATAGGAGCTG 21
    |||||
Qy 1328 GCCTTTGTGAAGCCAGCAGA 1347
    |||||
Db 20 GCCTTTGTGAAGCCAGCAGA 1
    |||||

```

Search completed: March 27, 2003, 07:47:14  
 Job time : 2780 secs

This Page Blank (uspto)

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:05:53 ; Search time 68 Seconds  
(without alignments)  
764.231 Million cell updates/sec

Title: US-09-744-125A-2

Perfect score: 1996  
Sequence: 1 NSARGVEPRRCSGSEPPSSS.....LINQDLIALIRAKKITPQ 390

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*  
1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1996	100.0	390	21	AA177554
2	1967	98.5	462	21	AA177555
3	1023	51.3	578	22	AB857874
4	160	8.0	465	21	AA177554
5	160	8.0	465	21	AA177554
6	160	8.0	479	21	AA177554
7	160	8.0	479	21	AA177554
8	160	8.0	490	21	AA177554
9	160	8.0	525	21	AA177554
10	123.5	6.2	108	22	AA177554

11	123.5	6.2	108	22	AA19989	Human liver associ
12	123.5	6.2	108	22	ABP40850	Human liver antige
13	116	5.8	2346	22	AB863519	Drosophila melanog
14	113.5	5.7	398	22	AB863199	Drosophila melanog
15	110.5	5.5	510	22	AB811764	Human LDL binding
16	110.5	5.5	510	22	AA179741	Human protein SEQ
17	110.5	5.5	530	19	AA179741	Human low density
18	110.5	5.5	530	22	AA179741	Human low density
19	110.5	5.5	546	22	AA179741	Human low density
20	110.5	5.5	546	22	AA179741	Human low density
21	109.5	5.5	739	22	AB862372	Human polypeptide
22	107	5.4	557	19	AA179741	Drosophila melanog
23	107	5.4	557	19	AA179741	Rabbit low density
24	106	5.3	519	23	AA179741	Lung small cell ca
25	106	5.3	897	21	AA179741	A rod shortened dy
26	106	5.3	900	21	AA179741	A rod shortened dy
27	106	5.3	1092	21	AA179741	A rod shortened dy
28	106	5.3	1201	21	AA179741	A rod shortened dy
29	106	5.3	1310	21	AA179741	A rod shortened dy
30	106	5.3	1310	21	AA179741	A rod shortened dy
31	106	5.3	3685	10	AA179741	Human Duchenne mus
32	106	5.3	3685	10	AA179741	Sequence encoded b
33	105.5	5.3	607	22	AB861226	Drosophila melanog
34	104.5	5.2	571	21	AA178794	Human antizua1-1 (
35	104.5	5.2	1219	21	AA178794	Human antizua1-2 (
36	104	5.2	367	21	AA178794	Nuclear transport
37	103	5.2	632	20	AA178794	B. burgdorferi ant
38	103	5.2	651	20	AA178794	B. burgdorferi ant
39	102.5	5.1	751	22	AB832409	Peptide #5060 enco
40	102.5	5.1	751	22	AB832409	Peptide #5368 enco
41	102.5	5.1	751	22	AB832409	Protein #5120 enco
42	102.5	5.1	751	22	AA178794	Human brain expres
43	102.5	5.1	751	22	AA178794	Human bone marrow
44	102.5	5.1	751	22	AA178794	Peptide #5068 enco
45	102.5	5.1	751	22	AA178794	Peptide #5305 enco

#### ALIGNMENTS

RESULT 1  
AA177554  
ID AA177554 standard; Protein, 390 AA.  
XX  
AC AA177554;  
XX  
DT 08-MAY-2000 (first entry)  
XX  
DE Human MIF1 protein (plasmid pCM480).  
XX  
MEK kinase; MEK1 interacting forkhead associated protein; MIF1; MEK2;  
XX  
FHA protein; forkhead associated protein; tumour; angiogenesis; human;  
XX  
psoriasis; cancer; anti-inflammatory; anti-asthmatic; immunosuppressant;  
XX  
vasotrophic; neuroprotective; antiarthritic; antiviral.  
XX  
Homo sapiens.  
XX  
W0200005362-A1.  
XX  
PD 03-FEB-2000.  
XX  
PP 21-JUL-1999; 99W0-EP05142.  
XX  
PR 21-JUL-1998; 98UG-0093590.  
XX  
PA (RHON ) RHONE-POULENC RORER SA.  
XX  
PI Marcieau C, Multon M, Polard-Housset V;  
XX  
DR WPI, 2000-195102/17.  
XX  
DR N-PSDB; AA258953.  
XX  
PT New MEK kinase interacting forkhead associated protein (MIF1) useful to

PT treat or diagnose, e.g. inflammation and tumors, and to identify its  
 PT specific modulators, to regulate MEK kinase activity -  
 Claim 20; Page 64-67; 78pp; English.

XX The invention provides MEK kinase (MEKK) interacting forkhead associated  
 CC (FHA) protein (MIF1). MIF1 is useful for screening for specific  
 CC modulators (potential therapeutic agents) and to reduce MEK activity in  
 CC cells. Antibodies specific to MIF1 are useful as diagnostic immunoassay  
 CC reagents to detect expression of MIF1, for purification of MIF1 and as  
 CC therapeutic (ant)agonists. The MIF1 nucleic acids are useful for:  
 CC recombinant production of MIF1, either in cultured cells or in vivo (gene  
 CC therapy); as source of probes and primers for detecting or quantifying  
 CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source  
 CC of therapeutic antisense sequences (used to increase MEK activity in  
 CC cells); and to identify inhibitors of MIF1. Regulation of MEK activity  
 CC via MIF1 regulation is useful for treatment of inflammation, asthma,  
 CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic  
 CC syndrome, neurodegeneration, tumors, angiogenesis, rheumatoid arthritis,  
 CC psoriasis or persistent viral infections. The present sequence represents  
 CC the MIF1 protein (plasmid pcM480).

XX Sequence 390 AA;

Query Match 100.0%; Score 1996; DB 21; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-171;  
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSARGVEPGRCSGSEPSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 60  
 DB 1 NSARGVEPGRCSGSEPSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 60  
 QY 61 GFWKRPANDLLINAVLQTNLDLTSVHLGVKFSCTFLREVOERWYALLVDPVSKLACQAM 120  
 DB 61 GFWKRPANDLLINAVLQTNLDLTSVHLGVKFSCTFLREVOERWYALLVDPVSKLACQAM 120  
 QY 61 GFWKRPANDLLINAVLQTNLDLTSVHLGVKFSCTFLREVOERWYALLVDPVSKLACQAM 120  
 DB 61 GFWKRPANDLLINAVLQTNLDLTSVHLGVKFSCTFLREVOERWYALLVDPVSKLACQAM 120  
 QY 121 ROLHPEALTAIOSKALFSAEBOQLSKVGSTQPTLETFODLHHRPDAFYLAFTAKALQ 180  
 DB 121 ROLHPEALTAIOSKALFSAEBOQLSKVGSTQPTLETFODLHHRPDAFYLAFTAKALQ 180  
 QY 181 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKDMDDEVLEHELMVADRQ 240  
 DB 181 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKDMDDEVLEHELMVADRQ 240  
 QY 241 KREIRQLEDELHKQVLDVSTIGMSSPDPDNOTLAVLGRMVRVYLMRSREITLGRATKDN 300  
 DB 241 KREIRQLEDELHKQVLDVSTIGMSSPDPDNOTLAVLGRMVRVYLMRSREITLGRATKDN 300  
 QY 301 QIDVDLSLEGPMKISRKQGVIKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 360  
 DB 301 QIDVDLSLEGPMKISRKQGVIKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 360  
 QY 361 VEIASLRFFVLINODLIALIRAEAAKITPQ 390  
 DB 361 VEIASLRFFVLINODLIALIRAEAAKITPQ 390

RESULT 2  
 AAY77555  
 ID AAY77555 standard; Protein; 462 AA.  
 XX  
 AC AAY77555;  
 DT 08-MAY-2000 (first entry)  
 XX  
 DE Human MIF1 protein (plasmid pcM577).  
 XX  
 KW MEK kinase; MEK interacting forkhead associated protein; MIF1; MEKK;  
 KW FHA protein; forkhead associated protein; tumor; angiogenesis; human;  
 KW psoriasis; cancer; anti-inflammatory; anti-asthmatic; immunosuppressant;  
 KW vasotropic; neuroprotective; antiarthritic; antiviral.  
 XX  
 OS Homo sapiens.

XX  
 PN WO200005362-A1.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 21-JUL-1999; 99WO-EP05142.  
 XX  
 PR 21-JUL-1998; 98US-0093590.  
 XX  
 PA (RHON) RHONE-POULENC RORER SA.  
 XX  
 PI Marcireau C, Multon M, Polard-Houssset V;  
 DR WPI: 2000-195102/17.  
 DR N-P8DB; AAZ58958.

PT New MEK kinase interacting forkhead associated protein (MIF1) useful to  
 PT treat or diagnose, e.g. inflammation and tumors, and to identify its  
 PT specific modulators, to regulate MEK kinase activity -  
 Claim 20; Page 70-73; 78pp; English.

XX The invention provides MEK kinase (MEKK) interacting forkhead associated  
 CC (FHA) protein (MIF1). MIF1 is useful for screening for specific  
 CC modulators (potential therapeutic agents) and to reduce MEK activity in  
 CC cells. Antibodies specific to MIF1 are useful as diagnostic immunoassay  
 CC reagents to detect expression of MIF1, for purification of MIF1 and as  
 CC therapeutic (ant)agonists. The MIF1 nucleic acids are useful for:  
 CC recombinant production of MIF1, either in cultured cells or in vivo (gene  
 CC therapy); as source of probes and primers for detecting or quantifying  
 CC genomic DNA encoding MIF1 or expression of mRNA encoding MIF1; as source  
 CC of therapeutic antisense sequences (used to increase MEK activity in  
 CC cells); and to identify inhibitors of MIF1. Regulation of MEK activity  
 CC via MIF1 regulation is useful for treatment of inflammation, asthma,  
 CC immunosuppression, cardiac ischemia or hypertrophy, myelodysplastic  
 CC syndrome, neurodegeneration, tumors, angiogenesis, rheumatoid arthritis,  
 CC psoriasis or persistent viral infections. The present sequence represents  
 CC the MIF1 protein (plasmid pcM577).

XX Sequence 462 AA;

Query Match 98.5%; Score 1967; DB 21; Length 462;  
 Best Local Similarity 99.0%; Pred. No. 2.2e-168;  
 Matches 386; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 3 AFG--GVBPGRCSGSEPSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 60  
 DB 73 AKGAGVEPGRCSGSEPSSEKKKVSAPSTVPSPAPAPGLTKRVKSKOPLQVTKDL 132  
 QY 61 GFWKRPANDLLINAVLQTNLDLTSVHLGVKFSCTFLREVOERWYALLVDPVSKLACQAM 120  
 DB 133 GFWKRPANDLLINAVLQTNLDLTSVHLGVKFSCTFLREVOERWYALLVDPVSKLACQAM 192  
 QY 121 ROLHPEALTAIOSKALFSAEBOQLSKVGSTQPTLETFODLHHRPDAFYLAFTAKALQ 180  
 DB 193 ROLHPEALTAIOSKALFSAEBOQLSKVGSTQPTLETFODLHHRPDAFYLAFTAKALQ 252  
 QY 181 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKDMDDEVLEHELMVADRQ 240  
 DB 253 AHMOLMKQYLLLEDQTVOPLPKGDQVLFNSDAEDLIDSKLKDMDDEVLEHELMVADRQ 312  
 QY 241 KREIRQLEDELHKQVLDVSTIGMSSPDPDNOTLAVLGRMVRVYLMRSREITLGRATKDN 300  
 DB 313 KREIRQLEDELHKQVLDVSTIGMSSPDPDNOTLAVLGRMVRVYLMRSREITLGRATKDN 372  
 QY 301 QIDVDLSLEGPMKISRKQGVIKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 360  
 DB 373 QIDVDLSLEGPMKISRKQGVIKLKNNGDFFLANEGRRPIYIDGRPVLCGSKMRLSNNSV 432  
 QY 361 VEIASLRFFVLINODLIALIRAEAAKITPQ 390  
 DB 433 VEIASLRFFVLINODLIALIRAEAAKITPQ 462

Query Match	51.3%;	Score 1023;	DB 22;	Length 578;
Best Local Similarity	55.5%;	Pred. NO. 3e-83;		
Matches 208;	Conservative	60;	Mismatches 97;	Indels 10;
Gaps	4;			
16	EPSSSEKKKVSAPSPVPPSPAPAPAGLTKRVKSKKOPL	-----QVTKDLAGKWPANDLLILIN	73	
204	KPPAHERSTSTSRSRHPVPAS	-----KKAQRGRGPRWGQVATYDGLGKWKIDDAALII	257	
74	AVLQTNLDLTVHVGVYKSCFTLRREVOERWVALLYDPVTSIKLACQAMROLHBEAIAAIOG	133		
258	GIQGTNDRLRIIHGKVGFSCKFTLQELQGWVALLLEPAVSRIVASAIRLHPDELVESVQR	317		
134	KALFSAKEQQLISKVGSISQPTLIEFTODLIHHNPDAFYLAARPAKALQAMWQMLKQYLLIE	193		
318	KALYSVQEBDLGLTIKSSQPKLEQFQELLIDKNASVFCARTAKSIQNNMILLKQYTLTP	377		
194	DOTVQPLPKGSD--QVNLNFSQAEPLIDDSKIKMDRDEVLEHEHLWAVADRQCRGRIROLEQELH	252		
378	DOSVAFPIYGTDOQPLSFSSAEDQIFPHDILNFRPDELAEHRKRLADRRNRGNRIILLENELS	437		
253	KWQVLVDSTI-TGWSBPDFDNQTLAVLRGSMVRYLMRSREITTLGRATKQNDIVDSLSEGP	311		

[illegible]

PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0138119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145219.  
PR 26-JUL-1999; 99US-0145276.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145931.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159337.  
PR 14-OCT-1999; 99US-0159633.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160988.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161933.  
PR 29-OCT-1999; 99US-0162142.

Query Match 8.0%; Score 160; DB 21; Length 465;  
Best Local Similarity 27.0%; Pred. No. 1,1e-05;  
Matches 47; Conservative 36; Mismatches 75; Indels 16; Gaps 3;

Oy	211	DADDLID-DSKAK-----DMEDEVLEHLMADRRCKEIRLDEELHKQVLVDSISG	263
Db	279	DDEEEDIDAMLRKTLVDDSDSCFNREEMWNMSKPHRALGLEO-----CTRTS	329
Oy	264	MSSPFDNQTLLAVLRGRWRYLRYSREITLGRATKONOIDVDLSLEGPAKMSRKOCVTK	323
Db	330	MORALMFHAIALVLRHPPDSKHPFKRKEVITLIGRSSGGLANDIDIDGKTKNYSKISRORALVK	389
Oy	334	LKNNGDFFIANEGRPRIYIDGRPVILGCSKRWLNNSVETIASLRFVFLINODLI	377
Db	390	LENYSFSLKNGKHQIILVNGGLDRGQIVTLLTSCSSINIRGITFVFKINKEAV	443
RESULT 5			
AAG40260			
ID	AAG40260 standard; Protein; 465 AA.		
XX	AAG40260;		
AC			
XX	18-OCT-2000 (first entry)		
DT			
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 49928.		
DE			
XX	Protein identification; signal transduction pathway; metabolic pathway;		
KM	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX			
OS	Arabidopsis thaliana.		
PN	EPI033405-A2.		
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	05-MAR-1999;	99US-0121825.	
PR	09-MAR-1999;	99US-0123180.	
PR	23-MAR-1999;	99US-0123548.	
PR	25-MAR-1999;	99US-0125788.	
PR	29-MAR-1999;	99US-0126264.	
PR	01-APR-1999;	99US-0126785.	
PR	06-APR-1999;	99US-0127462.	
PR	08-APR-1999;	99US-0128234.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	28-APR-1999;	99US-0130891.	
PR	30-APR-1999;	99US-0131449.	
PR	04-MAY-1999;	99US-0132407.	
PR	05-MAY-1999;	99US-0132484.	
PR	06-MAY-1999;	99US-0132485.	
PR	07-MAY-1999;	99US-0132486.	
PR	11-MAY-1999;	99US-0132487.	
PR	14-MAY-1999;	99US-0132863.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	20-MAY-1999;	99US-0134841.	
PR	21-MAY-1999;	99US-0135124.	
PR	24-MAY-1999;	99US-0135353.	
PR	25-MAY-1999;	99US-0135629.	
PR	27-MAY-1999;	99US-0136021.	
PR	28-MAY-1999;	99US-0136392.	
PR	01-JUN-1999;	99US-0136782.	
PR	03-JUN-1999;	99US-0137222.	
PR	04-JUN-1999;	99US-0137528.	
PR	05-AUG-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0133722.	
PR	08-JUN-1999;	99US-0138094.	
PR	10-JUN-1999;	99US-0138540.	
PR	14-JUN-1999;	99US-0138847.	
PR	16-JUN-1999;	99US-0139119.	
PR	17-JUN-1999;	99US-0139452.	
PR	18-JUN-1999;	99US-0139453.	
PR	18-JUN-1999;	99US-0139454.	
PR	18-JUN-1999;	99US-0139455.	
PR	18-JUN-1999;	99US-0139456.	
PR	18-JUN-1999;	99US-0139457.	
PR	18-JUN-1999;	99US-0139458.	
PR	18-JUN-1999;	99US-0139459.	
PR	18-JUN-1999;	99US-0139460.	
PR	18-JUN-1999;	99US-0139461.	
PR	18-JUN-1999;	99US-0139462.	
PR	18-JUN-1999;	99US-0139463.	
PR	18-JUN-1999;	99US-0139750.	
PR	18-JUN-1999;	99US-0139817.	
PR	21-JUN-1999;	99US-0139819.	
PR	22-JUN-1999;	99US-0139899.	
PR	23-JUN-1999;	99US-0140351.	
PR	24-JUN-1999;	99US-0140354.	
PR	28-JUN-1999;	99US-0140823.	
PR	29-JUN-1999;	99US-0140991.	
PR	30-JUN-1999;	99US-0141287.	
PR	01-JUL-1999;	99US-0142820.	
PR	01-JUL-1999;	99US-0142823.	
PR	02-JUL-1999;	99US-0142825.	
PR	06-JUL-1999;	99US-0142903.	
PR	08-JUL-1999;	99US-0143	

PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155639.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 8.0%; Score 160; DB 21; Length 465;  
Best Local Similarity 27.0%; Pred. No. 1, 1e-05;  
Matches 47; Conservative 36; Mismatches 75; Indels 16; Gaps 3;

QY 211 DAEDID-DSKUK-----DMRDEVLEHLMVADRKRREIRQLRQELHKMQVLVDSTIG 263  
Db 279 DGEERIDIDAMTRKLNLYVDDSDCFNRBEKNMSKHPHALGLBQ-----CIRTS 329  
QY 264 MSSPDNDQTLAVLRGRWRYLMRSREITLGRATKNOIDVDLSLEGPAWKISRQGVYK 323  
Db 330 MGRATMFHGAIAVLHCPDSKHFVRKREVIIGRSSGGLNVDIDLKYNVGSKISRQALVYK 389  
QY 324 LKNGDFFPLANEGRRPYITIDGRPVLCGSKWRLSNNSVVEIASLRVFLINDLI 377  
Db 390 LENYGSFSLKNGKOHILVNGKLDRCQIVTLTSCSINIRGITVFYFNKEAV 443  
  
RESULT 6  
AAG07417  
ID AAG07417 standard; Protein; 479 AA.  
XX AAG07417;  
XX 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4563.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 14-MAY-1999; 99US-0134376.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.



PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144326.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144633.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147915.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148355.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149358.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 16-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query March 8.0%; Score 160; DB 21; Length 479;  
Best Local Similarity 27.0%; Pred. No. 1.1e-05;

Matches	47;	Conservative	36;	Mismatches	75;	Indels	16;	Gaps	3
Oy	211	DAEDID-IDSKLK-----DMEDEVLEHLMVADRROKKEIRQLGEGLHKMOVLVDISTG	263						
Dd	293	DDEEIEDIDAMIRKLNVLVDDSDSCFNREEMNMSKPHRALIGLQ-----CTFRTS	343						
Oy	264	MSSPFDNQTALVLRGMRVRYLMRSREITLGRATKNOIDVDLSLEGPMKISRKGVIK	323						
Dd	344	MORAIMFHGALVLCPPDSKHFPRKREVIIGRSSGGLNVDDIDGKXNYGSKISRQALVK	403						
Oy	324	LKNNGDEPFANSGRPPIYIDGRPVLGSKMRLSNNSVVEIASLRFVFLINODLI	377						
Dd	404	LENYSFSLKNGKHQIHLVNGCKLRGQIVTLTSCSINIRGITTFVKINKEAV	457						
RESULT 7									
AAG40259									
ID	AAG40259 standard; Protein; 479 AA.								
XX	AAG40259;								
AC									
XX									
DT	18-OCT-2000 (first entry)								
XX									
De	Arabidopsis thaliana protein fragment SEQ ID NO: 49927.								
XX									
KW	Protein identification; signal transduction pathway; metabolic pathway;								
KM	hybridisation assay; genetic mapping; gene expression control; promoter;								
KX	termination sequence.								
XX									
Os	Arabidopsis thaliana.								
XX									
PN	EP1033405-A2.								
XX									
PD	06-SEP-2000.								
XX									
PF	25-FEB-2000; 2000EP-0301439.								
XX									
PR	25-FEB-1999;	99US-0121825.							
PR	05-MAR-1999;	99US-0123180.							
PR	09-MAR-1999;	99US-0123548.							
PR	23-MAR-1999;	99US-0125788.							
PR	25-MAR-1999;	99US-0126264.							
PR	29-MAR-1999;	99US-0126785.							
PR	01-APR-1999;	99US-0127462.							
PR	06-APR-1999;	99US-0128234.							
PR	08-APR-1999;	99US-0128714.							
PR	16-APR-1999;	99US-0129845.							
PR	19-APR-1999;	99US-0130077.							
PR	21-APR-1999;	99US-0130449.							
PR	23-APR-1999;	99US-0130510.							

PR	03-JUN-1999	99US-01375728
PR	04-JUN-1999	99US-01377202
PR	07-JUN-1999	99US-01377724
PR	08-JUN-1999	99US-01380094
PR	10-JUN-1999	99US-01385400
PR	10-JUN-1999	99US-01388447
PR	14-JUN-1999	99US-01391119
PR	16-JUN-1999	99US-01394552
PR	17-JUN-1999	99US-01394632
PR	18-JUN-1999	99US-01394544
PR	18-JUN-1999	99US-01394655
PR	18-JUN-1999	99US-01394576
PR	18-JUN-1999	99US-01394657
PR	18-JUN-1999	99US-01394663
PR	18-JUN-1999	99US-01397500
PR	18-JUN-1999	99US-01397653
PR	21-JUN-1999	99US-01398617
PR	21-JUN-1999	99US-01398699
PR	22-JUN-1999	99US-01400553
PR	23-JUN-1999	99US-01400554
PR	24-JUN-1999	99US-01400695
PR	28-JUN-1999	99US-01400823
PR	28-JUN-1999	99US-01400921
PR	30-JUN-1999	99US-01412877
PR	01-JUL-1999	99US-01418442
PR	01-JUL-1999	99US-01421554
PR	02-JUL-1999	99US-01420055
PR	06-JUL-1999	99US-01422900
PR	08-JUL-1999	99US-01428003
PR	09-JUL-1999	99US-01428200
PR	12-JUL-1999	99US-01428277
PR	13-JUL-1999	99US-01435244
PR	14-JUL-1999	99US-01440005
PR	15-JUL-1999	99US-01440085
PR	15-JUL-1999	99US-01440866
PR	16-JUL-1999	99US-01443425
PR	19-JUL-1999	99US-01443332
PR	19-JUL-1999	99US-01443332
PR	19-JUL-1999	99US-01443334
PR	19-JUL-1999	99US-01443334
PR	20-JUL-1999	99US-01443522
PR	20-JUL-1999	99US-01446332
PR	20-JUL-1999	99US-01446884
PR	21-JUL-1999	99US-01448114
PR	21-JUL-1999	99US-01450866
PR	21-JUL-1999	99US-01450888
PR	22-JUL-1999	99US-01450887
PR	22-JUL-1999	99US-01450887
PR	22-JUL-1999	99US-01450889
PR	22-JUL-1999	99US-01451922
PR	27-JUL-1999	99US-01455128
PR	27-JUL-1999	99US-01455139
PR	28-JUL-1999	99US-01455139
PR	28-JUL-1999	99US-01455139
PR	02-AUG-1999	99US-01463686
PR	02-AUG-1999	99US-01463689
PR	03-AUG-1999	99US-01470389
PR	04-AUG-1999	99US-01472704
PR	05-AUG-1999	99US-01471922
PR	05-AUG-1999	99US-01471922

PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151080.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 12-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161355.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 8.0%; Score 160; DB 21; Length 479;

Best Local Similarity 27.0%; Pred. No. 1,1e-05;  
Matches 47; Conservative 36; Mismatches 75; Indels 16; Gaps 3;  
Oy 211 DAEDLID-DSKLK-----DMRDEVLHEHLMVADNRQKREIREQEBLHKQVLVDSITG 263  
Db 293 DGEERIDIDAMIRKLNLPDSDSCFNEWMNSKPHHALGLEQ-----CIRTS 343  
Oy 264 MSSPDFDNOTLAVLRGRVYLRMSREITLGRATKDNQIDVDLSIEGPAWKISRQGVIK 323  
Db 344 MQRALMFHGAIAVLHCPDSKHFVRKREYITGRSSGGLAVDIDLKYNVGSKISRQALVK 403  
Oy 324 LKNNDFEFTIANEGRRPIYIDGRPVLCGSKMRLSNNSVVEIASLRFVLINDPLI 377  
Db 404 LENYGSFSLKMLGKHQIHLVNGKGLDRGQIVLTSCSSINIRGITVFVFINKEAV 457  
RESULT 8  
AAG07416  
ID AAG07416 standard; Protein; 490 AA.  
XX  
AC AAG07416;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4562.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
ED 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132484.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 14-MAY-1999; 99US-0134768.  
PR 18-MAY-1999; 99US-0134941.  
PR 19-MAY-1999; 99US-0135124.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140891.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0153363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155133.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 8.0%; Score 160; DB 21; Length 490;  
Best Local Similarity 27.0%; Pred.No.1,2e-05;  
Matches 47; Conservative 36; Mismatches 75; Indels 16; Gaps 3;

QY 211 DAEDLID-DSKLK-----DKRDEVLHEHLMWADRQREIROLEHKNQVLDSTG 263  
| | | | | : : : : | | | | : : : : | | | | : : : : | | | | : : : :  
DB 304 DGESEIDIDAMIRKLNLPVDDSDSCFNREEMNSGHPHALIGLEQ-----CTRTS 354  
| | | | | : : : : | | | | : : : : | | | | : : : : | | | | : : : :  
QY 264 MSSPDDFQTLAVLGRWVRYLMRGREITLGRATKQNDIVDLSLEGPAWKISRKQVIX 323  
| | | | | : : : : | | | | : : : : | | | | : : : : | | | | : : : :  
DB 355 MORALMFQALAVLHCPDSKHFVRKREVIIGRSSGGLAVDDLGKYNYSKISRQALVK 414  
| | | | | : : : : | | | | : : : : | | | | : : : : | | | | : : : :  
QY 324 LKNNGDFEIANEGRRPIYIDGRPYLCSKMRLSNNSVVEIASLRFVFLINDPLI 377  
| | | | | : : : : | | | | : : : : | | | | : : : : | | | | : : : :  
DB 415 LENVGSFELKVLKQKHLLVNGSKLDRGQIVTLTSCSSININGITVFINKKEAV 468  
| | | | | : : : : | | | | : : : : | | | | : : : : | | | | : : : :  
RESULT 9  
AAG40258  
ID AAG40258 standard; Protein; 525 AA.  
XX  
AC AAG40258;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49926.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.









PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI, 2001-457728/49.  
XX  
XX N-PSDB; AAS31700.  
XX  
XX Isolated nucleic acid molecule encoding a human liver related protein  
PT is used in preventing, treating or ameliorating disorders of the liver  
PT particularly cancer of the liver -  
XX  
XX Claim 11; SEQ ID No 176; 526pp; English.  
XX  
XX Sequences AAU19970-AAU20115 represent the liver associated polypeptides  
CC of the invention. Liver associated polypeptides and their associated  
CC polynucleotides are useful in the diagnosis, treatment and prevention of  
CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,  
CC cats, dogs, chickens or sheep. A pathological condition can be determined  
CC by detecting the presence or absence of a mutation in a liver associated  
CC polynucleotide. The treatable disorders include autoimmune diseases such  
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms  
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,  
CC cerebrovascular disorders such as cerebral ischaemia, nervous system  
CC disorders such as Alzheimer's disease, infections caused by bacteria,  
CC viruses and fungi, ocular disorders such as corneal infection, endocrine  
CC disorders such as premature labour and infertility, gastrointestinal  
CC disorders such as Crohn's disease, renal disorders such as  
CC glomerulonephritis and respiratory disorders such as asthma and pleurisy.  
CC The polypeptides can also be used to aid wound healing, to prevent skin  
CC aging due to sunburn, to maintain organs before transplantation, to  
CC regenerate tissues and in chemotaxis.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
XX Query Match 6.2%; Score 123.5; DB 22; Length 108;  
Best Local Similarity 71.1%; Pred. No. 0.0027;  
Matches 27; Conservative 3; Mismatches 5; Indels 3; Gaps 2;  
QY 315 ISRKQGVIKLKNNGDFTIANEGRRP-IYIDGRPVLCGS 351  
Db 64 LSASTGVIKLKNNGDFTIANEGRRPSTMDGR--VCGS 99  
RESULT 12  
ABP40850  
ID ABP40850 standard; Protein; 108 AA.  
XX  
XX AC ABP40850;  
XX  
XX 24-JUL-2002 (first entry)  
XX  
XX DE Human liver antigen HFLUF44, SEQ ID NO:176.  
XX  
XX Human liver antigen; liver disorder; hepatic disorder; infection;  
KM hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;  
KM cirrhosis; granulomatous hepatitis; toxin damage; drug damage;  
KM autoimmune disease; Wilson's disease; primary biliary cirrhosis;  
KM neoplastic disorder; cancer; tumour; portal hypertension;  
KM gastrointestinal disorder; hepatitis; drug screening; gene therapy;  
KM chromosome mapping; forensic analysis; antibody preparation;

KM hepatotropic; cytostatic; antiinflammatory; virucide; antibacterial;  
KM fungicide; parasiticide; antidote; immunosuppressive.  
XX  
XX OS Homo sapiens.  
XX  
XX US2002042096-A1.  
XX  
XX 11-APR-2002.  
XX  
XX 17-JAN-2001; 2001US-0764887.  
XX  
XX 31-JAN-2000; 2000US-179065P.  
XX  
XX 04-FEB-2000; 2000US-180628P.  
XX  
XX 28-JUN-2000; 2000US-214886P.  
XX  
XX 07-JUL-2000; 2000US-216647P.  
XX  
XX 07-JUL-2000; 2000US-216880P.  
XX  
XX 11-JUL-2000; 2000US-217487P.  
XX  
XX 11-JUL-2000; 2000US-217496P.  
XX  
XX 14-JUL-2000; 2000US-218290P.  
XX  
XX 26-JUL-2000; 2000US-220963P.  
XX  
XX 26-JUL-2000; 2000US-220964P.  
XX  
XX 14-AUG-2000; 2000US-224518P.  
XX  
XX 14-AUG-2000; 2000US-224519P.  
XX  
XX 14-AUG-2000; 2000US-225267P.  
XX  
XX 14-AUG-2000; 2000US-225268P.  
XX  
XX 14-AUG-2000; 2000US-225270P.  
XX  
XX 14-AUG-2000; 2000US-225447P.  
XX  
XX 14-AUG-2000; 2000US-225757P.  
XX  
XX 14-AUG-2000; 2000US-225758P.  
XX  
XX 22-AUG-2000; 2000US-226868P.  
XX  
XX 30-AUG-2000; 2000US-228924P.  
XX  
XX 01-SEP-2000; 2000US-229287P.  
XX  
XX 01-SEP-2000; 2000US-229343P.  
XX  
XX 01-SEP-2000; 2000US-229344P.  
XX  
XX 01-SEP-2000; 2000US-229345P.  
XX  
XX 05-SEP-2000; 2000US-229509P.  
XX  
XX 05-SEP-2000; 2000US-229513P.  
XX  
XX 08-SEP-2000; 2000US-231413P.  
XX  
XX 21-SEP-2000; 2000US-234223P.  
XX  
XX 21-SEP-2000; 2000US-234274P.  
XX  
XX 25-SEP-2000; 2000US-234997P.  
XX  
XX 27-SEP-2000; 2000US-235834P.  
XX  
XX 29-SEP-2000; 2000US-236327P.  
XX  
XX 29-SEP-2000; 2000US-236367P.  
XX  
XX 29-SEP-2000; 2000US-236368P.  
XX  
XX 29-SEP-2000; 2000US-236369P.  
XX  
XX 29-SEP-2000; 2000US-236370P.  
XX  
XX 02-OCT-2000; 2000US-236802P.  
XX  
XX 02-OCT-2000; 2000US-237037P.  
XX  
XX 02-OCT-2000; 2000US-237038P.  
XX  
XX 02-OCT-2000; 2000US-237039P.  
XX  
XX 02-OCT-2000; 2000US-237040P.  
XX  
XX 13-OCT-2000; 2000US-239935P.  
XX  
XX 20-OCT-2000; 2000US-240960P.  
XX  
XX 20-OCT-2000; 2000US-241785P.  
XX  
XX 20-OCT-2000; 2000US-241809P.  
XX  
XX 01-NOV-2000; 2000US-244617P.  
XX  
XX 17-NOV-2000; 2000US-249299P.  
XX  
XX 08-DEC-2000; 2000US-251856P.  
XX  
XX 08-DEC-2000; 2000US-251868P.  
XX  
XX 08-DEC-2000; 2000US-251869P.  
XX  
XX (ROSE/) ROSEN C A.  
XX  
XX (ROBE/) RUBEN S M.  
XX  
XX (BARA/) BARASH S C.  
XX  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI, 2002-381944/41.  
XX  
XX N-PSDB; ABN90055.  
XX  
XX New nucleic acid encoding human liver antigens, useful for diagnosis,  
PT treatment and prevention of e.g. hepatitis and hepatic cancer, also

PT related polypeptides and antibodies -  
XX Claim 11; SEQ ID No 176; 181pp; English.  
XX  
XX The invention relates to 145 novel human liver antigens (ABP40831-  
CC AB40975) and to cDNAs encoding them (ABN90036-ABN90180), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human liver antigen  
CC polypeptides, antibodies against human liver antigens, and the use of  
CC liver antigen polynucleotides and polypeptides in diagnosing, treating,  
CC prognosing or preventing various disorders of the liver. Such conditions  
CC include viral infections (e.g., cytomegalovirus, Epstein-Barr virus,  
CC hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic  
CC infections (e.g., Clonorchis sinensis, Echinococcus granulosus and  
CC Entamoeba histolytica), and also bacterial and fungal infections. Other  
CC disorders that may be treated include inflammatory conditions (e.g.,  
CC cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins,  
CC autoimmune diseases (e.g., Wilson's disease, primary biliary cirrhosis),  
CC neoplastic disorders (e.g., adenomas, haemangiomas and hepatocellular  
CC carcinoma), portal hypertension, or gastrointestinal disorders (e.g.,  
CC peptic ulcers, gastritis and peritoneal diseases). Liver antigen  
CC polypeptides and polynucleotides may also be used in screening for  
CC compounds which modulate liver antigen expression or activity. The  
CC polynucleotides may further be used for gene therapy, chromosome  
CC mapping, in the identification of individuals and in forensic analysis,  
CC and the polypeptides may be used as molecular weight markers or to  
CC prepare antibodies useful in disease diagnosis, drug targeting and  
CC phenotyping. The present sequence represents a human liver antigen of  
CC the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO at [seqdata.uspto.gov/sequence/](http://seqdata.uspto.gov/sequence/)  
CC  
XX  
SQ Sequence 108 AA;  
Query Match 6.2%; Score 123.5; DB 23; Length 108;  
Best Local Similarity 71.1%; Pred. No. 0.0027; Mismatches 5; Indels 3; Gaps 2;  
Matches 27; Conservative 3; Mismatches 5; Indels 3; Gaps 2;  
QY 315 ISRKQGVTKLNKNGDFFIANEGRRP-IYIDGRPVLCGS 351  
Db 64 LSASTGVTKLNKNGDFFIANEGRRPSTMDGR--VCGS 99  
RESULT 13  
ABB63519  
ID ABB63519 standard; Protein; 2346 AA.  
XX  
XX ABB63519;  
AC  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 17349.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
PF  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR  
XX  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PMD, Myers EW;  
PI  
XX

DR WPI; 2001-656860/75.  
DR N-PSDB; ABL07622.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
PS Disclosure; SEQ ID NO 17349; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (AB57737-AB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
CC  
XX  
SQ Sequence 2346 AA;  
Query Match 5.8%; Score 116; DB 22; Length 2346;  
Best Local Similarity 23.3%; Pred. No. 1;  
Matches 86; Conservative 55; Mismatches 126; Indels 102; Gaps 21;  
QY 45 KEVKSQKQLOVTKQGRKRPANDLLINAVIQTDLSVHGVFSGRFTIRVQEKMY 104  
Db 340 KRIQAEQELASANDL--LKQARESNLESATCQLAPSAV-----ASRLRSDLSLEL 391  
QY 105 ALVYQVTSKLCQAMROUHPAIAIOSKALFSAEBO--LISYVGSSTQPTLETPODL 162  
Db 392 YSMYAKSSEEL--EMRCEITQOL-KIQKSTIAIISAPILFQNDYQMKETNSL 447  
QY 163 LRRHPDAFY-----LARTAKAL-----QAHWQMKQYLLIEDQ----- 195  
Db 448 LREHDELLQNKLCLERELERALSTLNHNQENKQKQHTDLSRQVCMILDELNCIRAGV 507  
QY 196 ---TVQP---LPK-----GDQVNFSDAEDLID-DSKDKMDREVLHEHMAVADR----- 239  
Db 508 KHVRIQPTROLPTSSRLSDMLVTPSSIEELVDRNTYLLNMSREL--ELLEASERKNDK 565  
QY 240 -----QKREIRQLEQELHKWQVLSITGMSPPDPNQTLLAVLRGMVRY--LMSREIT 292  
Db 566 MLEQSKNHIKRLDA--RFALEBLLTQ-----KNTVTTLISKCRKYLTPAAQKX 616  
QY 293 LGRATKD-----NQIDVDSLSEGA--WKISRK-----QGVK---LKN 326  
Db 617 LGQNTVVDLDDSNLEPNDGALDTS--EQPANFESRKLEKRVARQLBQLEGEVKYASLKE 675  
QY 327 NGDFFIANE 335  
Db 676 NYDYVTSEK 684  
RESULT 14  
ABB63199  
ID ABB63199 standard; Protein; 398 AA.  
XX  
XX ABB63199;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 16389.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX

PD 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001MO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
PI Venter JC, Adams M, Li FWD, Myers EM;  
XX WPI, 2001-656860/75.  
DR N-PSDB; ABL07302.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX  
XX Disclosure; SEQ ID NO 16389; 21bp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB101840-AB16175), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (AB057737-AB072072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 398 AA;  
Query Match 5.7%; Score 113.5; DB 22; Length 398;  
Best Local Similarity 29.2%; Pred. No. 0.13;  
Matches 31; Conservative 20; Mismatches 34; Indels 21; Gaps 4;  
QY 2 SARGVPERGSCGSSPSSSEKKVSKAPSTVPSPAPAPGLTKRVKSKQPLQTKDGL 61  
Db 102 AAGASAPKAGCAGSTGKPS-----GNAPATP-PPPPPPPKAKSPVAKKPKLTKTFEY 154  
QY 62 RWKKN-----DL-----LLINAVLQTNLDTLSVHGVKFSR 93  
Db 155 RWRGDDQPGQTGYVDLEQCGAMVLDALIKIKNEMDPLTFRSQR 200  
RESULT 15  
AB11764  
ID AB11764 standard; peptide: 510 AA.  
XX  
XX ABB11764;  
XX  
XX 11-JAN-2002 (first entry)  
XX  
DE Human LDL binding protein homologue, SEQ ID NO:2134.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antiaesthetic; antiarthritic; haemostatic; antiarteriosclerotic;  
KW cytoskeletal; osteopathic; vasotropic; cardiac; vitruide; antibacterial;  
KW antifungal; vulnery; antitumor.  
XX  
XX Homo sapiens.  
XX

PN WO200157188-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 05-FEB-2001; 2001MO-US03800.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Dymanac RT;  
XX WPI, 2001-457740/49.  
DR N-PSDB; ABA09008.  
XX  
XX Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
XX e.g. arthritis and cancer -  
XX  
PS Claim 20; Page 240-241; 1963pp; English.  
XX  
XX Sequences AB10981-AB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a novel human  
CC polypeptide of the invention.  
XX  
SQ Sequence 510 AA;  
Query Match 5.5%; Score 110.5; DB 22; Length 510;  
Best Local Similarity 18.3%; Pred. No. 0.36;  
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;  
QY 1 NSARGVPERGSCGSSPSSSEKKVSKAPSTVPSPAPAG--LTKEVKS-- 50  
Db 102 NCGPGEAGAGFAEPDAKSTTYAARNSEPPPTVNGCKPSKDPTEIRGSD 161  
QY 51 -----KQPLQVTKDLGRWKPANDLLINAVLQ-TNDLTSVHGVKFSRFTLEVOERW 103  
Db 162 VGDRDHRPQKKKAGKGS-----ITLLMQTLNTLSTBEKTLAALCKYAEILLEHR 215

```

Qy 104 VALYDPIVSKLAQAMRQHLHEAIAIQAISKALPSKA----- 140
Db 216 NSQOMKLLQKKQSOLOVE--KDHIRGEHSKAVLARSKLESICRELQRHNRSLKEGVOR 273
Qy 141 --BEOLSKVSGSTSQPTLETFPQDILHRHPDAFYLAARTAKALQAHWQ-----LMKOY 190
Db 274 AREEEKKEVTSHFQVTLNDIQLOMEQHNE-----RNSKLNQENMELAEKLIQOYE 328
Qy 191 LLEDQTVQPLPKGDQVLFNSDAEDLIDSKLKMDEVLHEHLMVADRQKREIRQLEOE 250
Db 329 LREEH1-----DKVFKHKDLOQQLVDAKLQAOQEMLKE-----AERHQRE----- 369
Qy 251 LHKWQVIVDSITGMSSPDNDOTLAVLRGRWRYLMRSREI--TIGRATKQNDIVDL 308
Db 370 --KDFLLKEAVESQRMCELMKQOETHLKQOOLALYTEKEEEOFNTLSKS-----EFTTF 422
Qy 309 EGPWKISRKQGVIKLKNNGDFPIANEGRRPIYIDGRPVLGSKWRLSNNSVVEIASLRF 368
Db 423 KOEMEKMTKK--IKKLEKETTMY-----RSRWESSNKALLEMABEKT 462
Qy 369 V 369
Db 463 V 463

```

Search completed: March 27, 2003, 05:19:20  
 Job time : 75 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:18:00 ; Search time 25 Seconds  
(without alignments)  
458.998 Million cell updates/sec

Title: US-09-744-125A-2  
Perfect score: 1996  
Sequence: 1 NSARGVEPGRSGSEPPSS.....LINDLILIRAKITPQ 390

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued\_Patents\_AA.\*  
2: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110.5	5.5	530	4	US-08-979-608A-8
2	107	5.4	557	4	US-08-979-608A-5
3	102	5.1	10182	4	US-09-134-001C-3159
4	100	5.0	15281	2	US-08-471-119A-2
5	99.5	5.0	496	2	US-08-463-418-2
6	99.5	5.0	1377	2	US-08-308-818-4
7	98.5	4.9	855	2	US-08-468-558-2
8	98.5	4.9	855	2	US-08-676-444-2
9	97.5	4.9	1048	4	US-08-887-534A-85
10	97	4.9	1168	1	US-08-620-717A-9
11	97	4.9	1232	4	US-09-592-054-2
12	95.5	4.8	503	4	US-08-740-223A-18
13	95.5	4.8	503	4	US-09-202-491-4
14	95.5	4.8	503	4	US-09-202-491-10
15	95.5	4.8	503	4	US-09-709-188-18
16	92	4.6	731	2	US-08-313-185-51
17	92	4.6	731	2	US-08-458-499-15
18	92	4.6	731	3	US-09-082-614A-51
19	92	4.6	2101	1	US-08-466-390-4
20	92	4.6	2101	1	US-08-470-950-4
21	92	4.6	2101	1	US-08-467-781-4
22	92	4.6	2101	1	US-08-195-487-4
23	92	4.6	2101	2	US-08-483-924-4
24	92	4.6	2101	4	US-09-452-294-1
25	92	4.6	2101	5	PCT-US93-06160-4
26	91.5	4.6	1274	4	US-09-095-443-2
27	91	4.6	1167	1	US-08-485-568A-6

28	91	4.6	1167	2	US-08-590-554A-6	Sequence 6, Appli
29	91	4.6	1167	2	US-09-184-223-6	Sequence 6, Appli
30	90	4.5	303	1	US-08-185-432-5	Sequence 5, Appli
31	90	4.5	737	1	US-08-185-432-2	Sequence 2, Appli
32	90	4.5	737	1	US-08-185-432-4	Sequence 4, Appli
33	90	4.5	877	3	US-08-911-853-31	Sequence 31, Appli
34	90	4.5	877	4	US-09-479-409-31	Sequence 31, Appli
35	90	4.5	877	4	US-09-479-453-31	Sequence 31, Appli
36	89.5	4.5	442	1	US-08-391-615-6	Sequence 6, Appli
37	89.5	4.5	876	1	US-08-717-515-4	Sequence 4, Appli
38	89.5	4.5	1084	1	US-08-717-515-6	Sequence 6, Appli
39	89.5	4.5	1276	1	US-08-717-515-8	Sequence 8, Appli
40	89.5	4.5	1432	3	US-08-781-891-71	Sequence 71, Appli
41	89	4.5	555	4	US-09-813-872-4	Sequence 4, Appli
42	89	4.5	1137	1	US-08-369-043-2	Sequence 2, Appli
43	89	4.5	1193	4	US-09-237-725A-4	Sequence 4, Appli
44	89	4.5	2154	2	US-08-841-349-4	Sequence 4, Appli
45	88.5	4.4	495	4	US-09-351-457-5	Sequence 5, Appli

#### ALIGNMENTS

RESULT 1  
US-08-979-608A-8  
Sequence 8, Application US/08979608A  
Patent No. 6355451  
GENERAL INFORMATION:  
APPLICANT: Lees, Ann M.  
Lees, Robert S.  
Law, Simon W.  
Arjona, Anibal A.  
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fleish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979, 608A  
FILING DATE: 26-NOV-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/048, 547  
FILING DATE: 03-JUN-1997  
APPLICATION NUMBER: US 60/031, 930  
FILING DATE: 27-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: 10797-002001 (Formerly 3963/59818)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-979-608A-8  
Query Match 5.5%; Score 110.5; DB 4; Length 530;

Best Local Similarity 18.3%; Pred. No. 0.063;  
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVGPGRSGSPSSSEKKVSKAPSTVPSPSPA-----PAPG--LTKRVKKS-- 50  
DB 75 NCGGDEDAQGEPAPEAEKSRVTYVANGPEPEFPVYGEKEPEKGDPTNEIRQSD 134  
QY 51 -----KQPLQVTKDGRKKRANLILINAVIQT--NDLTSVHLGVKFSQRTLRVQERW 103  
DB 135 VGDHRHRRPQEKKAKGGLKE-----ITLMQTLNTLSTPEEKLAALCKKYAEILLEHR 188  
QY 104 YALLVDYPSKLAQACQMLHPEALAIQSKALFSA----- 140  
DB 189 NSQKMKLIQKQKQSLVQE--KDLHGEHSKAVLARSKLESICRELQRRNRSLKEGVOR 246  
QY 141 ---EEOQLSKVSTSQPTLETPODLLHRHPDAFYLTARTAKALOAHWQ-----LMKQYV 190  
DB 247 ABEERKREKREVTSHFQVTLNDIQLOMEQHNE-----RNSKLQENNELAERLKLIEQYE 301  
QY 191 LLEDQTVQPLPKGDQVLYNSDAEDLIDSKLDMRDEVLEHELMVADRQKREIRQLEGE 250  
DB 302 LREEH1-----DKVFKKKDLQOQLVDAKLOAQOEMLKE-----AEEHORE----- 342  
QY 251 LHKQVLYVDSITGSSPDNDQTLAVLRGRWRYLMRSREI--TLGRATKQNDIVDSL 308  
DB 343 --KDFLKEAVSQRMCMLKQOETHLKOQALALYTEKEFEFQNTLSKS-----EVFTTF 395  
QY 309 EGPWAKIRKQGVYKLNKNGDFIANEGRRPIYIDGRPVLCGSKWELSNNSVEIASLRF 368  
DB 396 KQEMEMTKK--IKKLEKETMY-----RSRWESSKALLLEMAEKT 435  
QY 369 V 369  
DB 436 V 436

## RESULT 2

US-08-979-608A-5  
; Sequence 5, Application US/08979608A  
; Patent No. 6355451

## GENERAL INFORMATION:

APPLICANT: Lees, Robert S.  
Law, Simon W.  
Aljona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA

COUNTRY: USA  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: PASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,608A  
FILING DATE: 26-NOV-1997  
PRIORITY DATE: 6355451-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/048,547  
FILING DATE: 03-JUN-1997

APPLICATION NUMBER: US 60/031,930  
FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-08-979-608A-5

Query Match 5.4%; Score 107; DB 4; Length 557;

Best Local Similarity 18.9%; Pred. No. 0.14; Indels 114; Gaps 17;

Matches 76; Conservative 76; Mismatches 137;

QY 16 EPPSSSEKKVSKA-----PSTPV---PSPAPAPGLTKRVKKS-----KQPLQVT 57  
DB 105 EPEDAEKSAVYARNGPEPEPGTVPVNGEKEYSKAPG--TEIIRTSDEVDGRHRRPQEK 163  
QY 58 KDLGRWKPANDLILINAVIQT--NDLTSVHLGVKFSQRTLRVQERWYALLYDPIYSKLA 116  
DB 164 KAKGLKE-----ITLMQTLNTLSTPEEKLAALCKKYAEILLEHRNSQKMKLIQKQ 217  
QY 117 QQAMQHLHEALAIQSKALFSA----- 150  
DB 218 SOLVQE--KDLHGEHSKAVLARSKLESICRELQRRNRSLKEGVQARABEERKREKVT 275  
QY 151 TSQPTLETPODLLHRHPDAFYLTARTAKALOAHWQ-----LMKQYLLDQTVQPLPKG 203  
DB 276 HFQMTLNDIQLOMEQHNE-----RNSKLQENNELAERLKLIEQYELREH1----- 323  
QY 204 DQVLYNSDAEDLIDSKLDMRDEVLEHELMVADRQKREIRQLEGEHLHKQVLYVDSITG 263  
DB 324 DKVFKKKDLQOQLVDAKLOAQOEMLKE-----AEEHORE-----KQPLLEAVES 369  
QY 264 MSSPDNDQTLAVLRGRWRYLMRSREI--TLGRATKQNDIVDSLSEGPWAKIRKQGV 321  
DB 370 QRMCELMKQOETHLKOQALALYTEKEFEFQNTLSKS-----EVFTTFQEMEMTKK--I 422  
QY 322 IKLNKNGDFIANEGRRPIYIDGRPVLCGSKWELSNNSVEIASLRF 364  
DB 423 KKLKLEKETMY-----RSRWESSKALLLEMAEKT 447

## RESULT 3

US-09-134-001C-3159  
; Sequence 3159, Application US/09134001C  
; Patent No. 6380370

## GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NOCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 3159

LENGTH: 10182

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3159

Query Match 5.1%; Score 102; DB 4; Length 10182;

Best Local Similarity 24.3%; Pred. No. 28;

Matches 83; Conservative 60; Mismatches 136; Indels 62; Gaps 21;

QY 16 EPPSSSEKKK-----VSKAPSTVPSPAPAPGLTKR--VKSKQPLQVTKDGRWKPANDLIL 70

Db 3424 EPLETAKQLOANDQKST-----DGMTQOSVQSYERKLOEAKD-----KIN 3466  
Qy 71 LINVLQON-DLTSVHLGVKSCRTFLREVQERWALLYD--PVL-SKLACQAMROLHPE 126  
Db 3467 SINVLANNPVNARLRN-KVETEQINNELTQAKGLTVQKPLINAKTALQOOSIDNOPS 3525  
Qy 127 AI-----AAIOS-KALFSAEE--QLSKVGSSTQPLETFODLIHRHPDAFLARTAK-A 178  
Db 3526 TTGMTAETIQVNNARQAEQVIONANKIENAOBSVQOVSEKSKVQDALSELNNAKSA 3585  
Qy 179 LOAHWQMKQYLLLEDQVQPL-----PKGDVLFNSDAEDLIDDSKDKMREDEVL-EH 231  
Db 3586 LRADKQELQAV---NQLIQPTDLNNKKPASITAYNQRVQCFNSNELNSTKTITDRILKEQ 3642  
Qy 232 ELMWAD-RQKREIRQLEQELHKQVLYDSITGMSSPDPNQTLAVLGRMVRVY--MRS 288  
Db 3643 NPVSADVNNALNKREAVQCKNEARALLQON-----KEDNSALVRAREQLQAVDVQPS 3695  
Qy 289 REITLGRATKDNQIDVDLSLEGPAWKISRQGVITLKNNGD 329  
Db 3696 TE-GMTQOTKD---DYNKQQAQOEISKAQVY---DNGD 3729

## RESULT 4

US-08-471-119A-2  
; Sequence 2, Application US/08471119A  
; Patent No. 5827706  
; GENERAL INFORMATION:  
; APPLICANT: Leitner, Ernst  
; APPLICANT: Schneider, Elisabeth  
; APPLICANT: Schoerendorfer, Kurt  
; APPLICANT: Weber, Gerhard  
; TITLE OF INVENTION: Cyclosporin Synthetase  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: No. 5827706artis Corporation  
; STREET: 59 Route 10  
; CITY: East Hanover  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07936  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471.119A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kassenoft, Melvyn  
; REGISTRATION NUMBER: 26,389  
; REFERENCE/DOCKET NUMBER: 100-8029/CONT/CONT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 503 8474  
; TELEFAX: 201 503 8807  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15281 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHEITICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Tolypocladium niveum  
; STRAIN: ATCC 34921  
; US-08-471-119A-2

Query Match 5.0%; Score 100; DB 2; Length 15281;

Best Local Similarly 23.2%; Pred. No. 78;  
Matches 73; Conservative 47; Mismatches 118; Indels 76; Gaps 14;  
Qy 14 GSESSSEKKKVSAPSPVPSPAPAPGLTKRVKSKQVQLVTKDGRMKPAN-DILLI 72  
Db 2126 GLEPSQSVAQPVNKAQV-----PFGLE--GKQVHGVTAMDRTGLSALSPDLIVI 2174  
Qy 73 NAVIQTNLDLTVHGVKSCRTFLREVQERWALLYDPVLSKLACQAMRQ-----LHP 125  
Db 2175 NSVAQ-----YPSREYLAIVE---ALVRIDGVARIFPGDMRTYATHKDFLVA 2220  
Qy 126 EALAIQSKALFSAEQELSKVGSSTQPLETFODLIHRHPDAFYLTAKALQAHWOL 185  
Db 2221 RAVHTNGSKVTRSKVQGV-----ARLEIEEELLVDP-AFFTS----- 2258  
Qy 186 MKQYLLLEDQVQPLPKGDVLFNSDAEDLIDDSKDKMREDEVLEHELMVADRKQREIR 245  
Db 2259 LKESISEEIEHEVILPKMKKNN-----ELSSRYGAVLH-IRHNQNSRSIH 2306  
Qy 246 QLEQELHKM-----QVLVDSITGMSSPDPNQTLAVLGRMVRVYLRREITLGRATKD 299  
Db 2307 KINAE--SWIDFASQMDRQGLARLKENKDAESIAVNIYKTIIV-ERHIAKSLADH 2363  
Qy 300 NQIDVDLSLEGPAW 313  
Db 2364 DDDPTSSIDGVAW 2377

## RESULT 5

US-08-463-418-2  
; Sequence 2, Application US/08463418  
; Patent No. 5908971  
; GENERAL INFORMATION:  
; APPLICANT: Van Der Straeten, Dominique et al.  
; TITLE OF INVENTION: CRUCIFER ACC SYNTHASE AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463.418  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/962,481  
; FILING DATE: 15-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/161002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 496 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-463-418-2

Query Match 5.0%; Score 99.5; DB 2; Length 496;

Best Local Similarity 20.2%; Pred. No. 0.6; Matches 68; Conservative 56; Mismatches 117; Indels 95; Gaps 12;

```

Qy 32 PVPSPAPAPGLTK-----RVKSKQPLQVTDLGRKWPANDLLINAVLQNTDLSVHL 86
Db 170 PVPSSSDNFKLTVADEAWYKKAQSNKKVGLITNPSNPL---GTMLDKDTLTNL-- 224
Qy 87 GYKESCRFTLRVQERWYLLY----DPVISKACQAMQRLPEAIAIOSKALPSKAE 142
Db 225 -VRFVTRKNHILVDEIYATVPAGDFV-----SVAEVDNDVISEVNV 268
Qy 143 QLSKVGSTSP-----TLTFQDL---HRHPDAFYLTAKALQAHQMLKQY 190
Db 269 DLHIVYSISKMGJPGFVGIYSDNSVSCARMSFGL-----VSSQTQLMAM 322
Qy 191 ILEDQTVQPLPKGDVLFNSDAEDLIDSKLMDRDEVLHEHLMVAD----- 237
Db 323 LSDQGFVD-----NF-----LMSSRRIGIRHKVFTTGKKADIACLTNSAGLFAM 368
Qy 238 ---RQKREIRQLEQELHMQVLVDSITGMSF-----DFDQTLAVL 277
Db 369 MDRLHLRRNSFESIELMHTIIDVKLVSPGSSFRCTEPGWFRCFANMDDTLHVA 428
Qy 278 RGRMYRLMRSREITLGRATKDNQIDVDLSLGPAP 313
Db 429 IGRIDFVSKNKKIYKASNDQVQNSAKKTKM 464

```

RESULT 6  
US-08-308-818-4

; Sequence 4, Application US/08308818  
; Patent No. 5847077

; GENERAL INFORMATION:  
; APPLICANT: Green, Michael R

; APPLICANT: Reese, Joseph C

; TITLE OF INVENTION: A No. 5847077el Fungal Multisubunit Protein

; TITLE OF INVENTION: Complex Critical for Expression of Fungal Proteins

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby PC

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: US

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/308, 818

; FILING DATE: 19-SEP-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Ludwig, S. Peter

; REGISTRATION NUMBER: 25,351

; REFERENCE/DOCKET NUMBER: 0342/0A404

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-52707700

; TELEPHONE: 212-753-6237

; TELEFAX: 236687

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1377 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: D. melanogaster

; IMMEDIATE SOURCE:

; CLONE: TAF11-250  
US-08-308-818-4

Query Match 5.0%; Score 99.5; DB 2; Length 1377;  
Best Local Similarity 21.5%; Pred. No. 2.6;

Matches 68; Conservative 51; Mismatches 100; Indels 97; Gaps 15;

```

Qy 5 GGVEPGRCS-----GSEPSSEKKVSKAPSPVPPSP-----PAPGLTKVSKSK 51
Db 998 GPADPTGGEFGSYRVNPKPTQKEQESQ-PKRSVGTIDBLRLPLQPAKELLROPK 1056
Qy 52 QPLQVTKDLGRKWPANDLLINAVLQNTDLSVHLGV-KFS--CEFTLRVQERWYLLY 108
Db 1057 VVEEIKKLSRY---EVIDVRTLSTEKAKGEGGMKFSNGNFSIAHQERY----- 1107
Qy 109 DPVISKACQAMQRLPEAIAIQ---SKALFSAEQLSKVG----- 149
Db 1108 ---KECQRIFDLQNVLASSEVLSTDEASSASBSDELQKNENMLSNKKTSTQ 1162
Qy 150 -STSQPTLETQDL-----HRHPDAFYLTAKALQAHQML----- 186
Db 1163 USRESEELER-QLRLQDEEHGPGSSGAGAKGKDDPGQMLATNNQRIILRTTF 1221
Qy 187 -----KOYLLDQTVQPLPKGDVLFNSDAEDLIDSKLMDRDEVLHEHLMVADRSQ- 240
Db 1222 RNDQKEVTRVE--TVRQPVIDAVI-----KIRTKQEQFIKQFATLDEQOK 1267
Qy 241 ---KREIRQLEQELH 253
Db 1268 EEMKREKRRIOELRR 1283

```

RESULT 7  
US-08-468-558-2

; Sequence 2, Application US/08468558

; Patent No. 5877280

; GENERAL INFORMATION:

; APPLICANT: Wetmur, James G.

; TITLE OF INVENTION: Cloning and Expression of Thermostable

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: United States of America

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468, 558

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: MSM94-05

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 855 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-468-558-2

Query Match 4.9%; Score 98.5; DB 2; Length 855;







```
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/022/999
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cobert, Robert J
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 333
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: TIE ligand-4
LOCATION: 1..503
OTHER INFORMATION:
US-08-740-223A-18

Query Match
Best Local Similarity 21.8%; Score 95.5; DB 4; Length 503;
Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;

QY 30 STVPSPAPAPGLTKRVKKS-KOPLQVTKDLGRWKPANDLLINAVLQTNL-LTSVHL 86
D 52 SEPCPGEVSRDNTLQRESLANPLHIGK-----LPTQVQKQLEALQNTQWMLKLER 106
QY 87 GVKSCRTLEVEQERWALLYDPVISKLACQAMQRLPEALIAIQSKALFSKAEQILLS 146
D 107 AIKTLRSKLEQVOQMAQNTAPML-ELGTSILNQ-----TTAQIRK---LTDMEAQLLN 158
QY 147 KVGSTQPTLETF-----QDLHRHPDAFYARTAKALQAHQMLM--KQYLLLEDQTV 197
D 159 QTSRMDAQMPETFLSTKNLEQOLLQROKLOLOGQNSALEKRLQALETKQ---QEELA 214
QY 198 QPLPKGDQVINFSDAEDLIDSKLDMRDEVLEHELMVADRQRKEIRQLQELHKQVYL 257
D 215 SILSKAKATLNTLSRQSAALNTNIEGLRG-VRHNSLSLQD--QCHSLROL--LVLLRHL 268
QY 258 VDSITGSSPFDNQTLAVLGRWV---RYLMS-----REITLGATKQNOIDVDL 306
D 269 VOERANASAPAF-----IMAGEQVFQDCAEIQRSQASAGVYTIQVSNATPRKVFCDL 322
QY 307 SLEGPAMKI--SRKQGVILKKN-----GDPFIANE-----GRPIY 341
D 323 OSSGGRMTLQIRENGTVNFORMKDYQKGRGDPAGEHMLGNEVVHQLTRRAY 376

RESULT 13
US-09-202-491-4
Sequence 4, Application US/09202491
Patent No. 6432667
GENERAL INFORMATION:
APPLICANT: Valenzuela et al.
TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
FILE REFERENCE: REG330-K
CURRENT APPLICATION NUMBER: US/09/202,491
EARLIER FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: PCT/US97/10728
EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: 60/022,999
EARLIER FILING DATE: 1996-08-02
EARLIER APPLICATION NUMBER: 60/021,087
EARLIER FILING DATE: 1996-07-02
EARLIER APPLICATION NUMBER: 08/665,926
EARLIER FILING DATE: 1996-06-19
NUMBER OF SEQ ID NOS: 14
```

```
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 503
TYPE: PRT
ORGANISM: Homo sapiens
US-09-202-491-4

Query Match
Best Local Similarity 21.8%; Score 95.5; DB 4; Length 503;
Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;

QY 30 STVPSPAPAPGLTKRVKKS-KOPLQVTKDLGRWKPANDLLINAVLQTNL-LTSVHL 86
D 52 SEPCPGEVSRDNTLQRESLANPLHIGK-----LPTQVQKQLEALQNTQWMLKLER 106
QY 87 GVKSCRTLEVEQERWALLYDPVISKLACQAMQRLPEALIAIQSKALFSKAEQILLS 146
D 107 AIKTLRSKLEQVOQMAQNTAPML-ELGTSILNQ-----TTAQIRK---LTDMEAQLLN 158
QY 147 KVGSTQPTLETF-----QDLHRHPDAFYARTAKALQAHQMLM--KQYLLLEDQTV 197
D 159 QTSRMDAQMPETFLSTKNLEQOLLQROKLOLOGQNSALEKRLQALETKQ---QEELA 214
QY 198 QPLPKGDQVINFSDAEDLIDSKLDMRDEVLEHELMVADRQRKEIRQLQELHKQVYL 257
D 215 SILSKAKATLNTLSRQSAALNTNIEGLRG-VRHNSLSLQD--QCHSLROL--LVLLRHL 268
QY 258 VDSITGSSPFDNQTLAVLGRWV---RYLMS-----REITLGATKQNOIDVDL 306
D 269 VOERANASAPAF-----IMAGEQVFQDCAEIQRSQASAGVYTIQVSNATPRKVFCDL 322
QY 307 SLEGPAMKI--SRKQGVILKKN-----GDPFIANE-----GRPIY 341
D 323 OSSGGRMTLQIRENGTVNFORMKDYQKGRGDPAGEHMLGNEVVHQLTRRAY 376

RESULT 14
US-09-202-491-10
Sequence 10, Application US/09202491
Patent No. 6432667
GENERAL INFORMATION:
APPLICANT: Valenzuela et al.
TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
FILE REFERENCE: REG330-K
CURRENT APPLICATION NUMBER: US/09/202,491
EARLIER FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: PCT/US97/10728
EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: 60/022,999
EARLIER FILING DATE: 1996-08-02
EARLIER APPLICATION NUMBER: 60/021,087
EARLIER FILING DATE: 1996-07-02
EARLIER APPLICATION NUMBER: 08/665,926
EARLIER FILING DATE: 1996-06-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 10
TYPE: PRT
ORGANISM: Homo sapiens
US-09-202-491-10

Query Match
Best Local Similarity 21.8%; Score 95.5; DB 4; Length 503;
Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;

QY 30 STVPSPAPAPGLTKRVKKS-KOPLQVTKDLGRWKPANDLLINAVLQTNL-LTSVHL 86
D 52 SEPCPGEVSRDNTLQRESLANPLHIGK-----LPTQVQKQLEALQNTQWMLKLER 106
QY 87 GVKSCRTLEVEQERWALLYDPVISKLACQAMQRLPEALIAIQSKALFSKAEQILLS 146
D 107 AIKTLRSKLEQVOQMAQNTAPML-ELGTSILNQ-----TTAQIRK---LTDMEAQLLN 158
```

```
OY 147 KVGSTSQPTLETF-----QDLHRHPDAFYIARTAKALOAHWOLM--KQYYLEEDQTV 197
      :      |||      :      :      :      :      :      :      :
DB 159 QTSRMDAQMPETFLSTNKLLENQLLQROKLOQLOQONSALERKLOLETQO---QEELA 214
      :      :      :      :      :      :      :      :      :
OY 198 QLPKQDQVLNFSDAEDLIDDSKDKMDEVELEHMLVADRQKREIROLEQELHKQVUL 257
      :      :      :      :      :      :      :      :      :
DB 215 SLISKAKALNLTLSKOSALTNIERGLRG-VRHNSSLQD--QOHSRLQD---LVILRHL 268
      :      :      :      :      :      :      :      :      :
OY 258 VDSITGMSSPDFDNOQLAVLRGRMV---RYLMRS-----REITLGRATKDNQIDVDL 306
      :      :      :      :      :      :      :      :      :
DB 269 VOERANASAPAF-----IMAGEQVFQDCAEIQRSASASGVYTIQVSNATKPRKVFCDL 322
      :      :      :      :      :      :      :      :      :
OY 307 SLEGPAMKI--SRKQGVILKNN-----GDFFIANE-----GRPIY 341
      :      :      :      :      :      :      :      :      :
DB 323 OSSGGRWTLIQRRNGTVNFORWKDYKQGFQDPAGEHMLGNEVNHQLTRAAAY 376
      :      :      :      :      :      :      :      :      :
```

## RESULT 15

```
US-09-709-188-18
; Sequence 18, Application US/09709188
; Patent No. 6441137
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed ligand - Vascular Intercellular Signaling Molecule
; FILE REFERENCE: REG 333-Z
; CURRENT APPLICATION NUMBER: US/09/709,188
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-709-188-18
```

```
Query Match 4.8%; Score 95.5; DB 4; Length 503;
Best Local Similarity 21.8%; Pred. No. 1.4;
Matches 77; Conservative 51; Mismatches 155; Indels 71; Gaps 18;
```

```
OY 30 STVPSPSPAPAPGLTKRVKKS--KQPLQVTKDLGRWKRPANDLLINAVLOTND--LTSVHL 86
      :      :      :      :      :      :      :      :      :
DB 52 SEBCPPGPEVSRDSNTLQRESLANPLHLGK---LPTQOVKQLEQALQNTQWMLKKER 106
      :      :      :      :      :      :      :      :      :
OY 87 GYKFSQFTLREYQERWYALLYPVISTKLACQAMRQLHPBAIAIOSKALFSKAEBOILS 146
      :      :      :      :      :      :      :      :      :
DB 107 AKITLRSLKEQVQOQMAQNOTAPML-ELGTSILNQ---TTAQIRK---LTDMEHQLLN 158
      :      :      :      :      :      :      :      :      :
OY 147 KVGSTSQPTLETF-----QDLHRHPDAFYIARTAKALOAHWOLM--KQYYLEEDQTV 197
      :      :      :      :      :      :      :      :      :
DB 159 QTSRMDAQMPETFLSTNKLLENQLLQROKLOQLOQONSALERKLOLETQO---QEELA 214
      :      :      :      :      :      :      :      :      :
OY 198 QLPKQDQVLNFSDAEDLIDDSKDKMDEVELEHMLVADRQKREIROLEQELHKQVUL 257
      :      :      :      :      :      :      :      :      :
DB 215 SLISKAKALNLTLSKOSALTNIERGLRG-VRHNSSLQD--QOHSRLQD---LVILRHL 268
      :      :      :      :      :      :      :      :      :
OY 258 VDSITGMSSPDFDNOQLAVLRGRMV---RYLMRS-----REITLGRATKDNQIDVDL 306
      :      :      :      :      :      :      :      :      :
DB 269 VOERANASAPAF-----IMAGEQVFQDCAEIQRSASASGVYTIQVSNATKPRKVFCDL 322
      :      :      :      :      :      :      :      :      :
OY 307 SLEGPAMKI--SRKQGVILKNN-----GDFFIANE-----GRPIY 341
      :      :      :      :      :      :      :      :      :
DB 323 OSSGGRWTLIQRRNGTVNFORWKDYKQGFQDPAGEHMLGNEVNHQLTRAAAY 376
      :      :      :      :      :      :      :      :      :
```

Search completed: March 27, 2003, 05:23:22  
Job time : 40 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:21:49 ; Search time 30 Seconds  
(without alignments)  
763.408 Million cell updates/sec

Title: US-09-744-125a-2

Perfect score: 1996  
Sequence: 1 NSARGVEPGRCSGSEPS...LINDLIATRAAKITPQ 390

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCr\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCrUS\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123.5	6.2	108	10 US-09-764-887-176	Sequence 176, App
2	110.5	5.5	530	9 US-09-976-740-8	Sequence 8, Appl
3	110.5	5.5	530	10 US-09-962-055-8	Sequence 8, Appl
4	110.5	5.5	530	12 US-10-023-529-8	Sequence 8, Appl
5	110.5	5.5	530	12 US-10-023-523-8	Sequence 8, Appl
6	110.5	5.5	546	9 US-09-976-740-44	Sequence 44, Appl
7	110.5	5.5	546	12 US-10-023-529-44	Sequence 44, Appl
8	110.5	5.5	546	12 US-10-023-523-44	Sequence 44, Appl
9	107	5.4	557	9 US-09-976-740-5	Sequence 5, Appl
10	107	5.4	557	10 US-09-962-055-5	Sequence 5, Appl
11	107	5.4	557	12 US-10-023-529-5	Sequence 5, Appl
12	107	5.4	557	12 US-10-023-523-5	Sequence 5, Appl
13	106	5.3	519	10 US-09-833-790-435	Sequence 435, App
14	102.5	5.1	751	10 US-09-864-761-38419	Sequence 38419, A
15	100.5	5.0	475	10 US-09-815-242-11259	Sequence 11259, A
16	99	5.0	327	10 US-09-777-745-4	Sequence 4, Appl
17	97.5	4.9	1048	10 US-09-741-669-409	Sequence 409, App
18	97.5	4.9	1048	10 US-09-815-242-10062	Sequence 10062, A
19	96	4.8	558	10 US-09-777-745-6	Sequence 6, Appl

20	94.5	4.7	465	9 US-10-102-806-676	Sequence 676, App
21	94.5	4.7	1033	9 US-09-820-843a-75	Sequence 75, Appl
22	93.5	4.7	277	10 US-09-925-300-1467	Sequence 1467, App
23	93	4.7	302	9 US-10-043-487-248	Sequence 248, App
24	92.5	4.6	1312	10 US-09-737-178-120	Sequence 120, App
25	92	4.6	712	10 US-09-925-302-678	Sequence 678, App
26	91.5	4.6	526	10 US-09-777-745-2	Sequence 2, Appl
27	91.5	4.6	1274	9 US-10-020-215-2	Sequence 327, App
28	90.5	4.5	573	9 US-10-043-487-327	Sequence 3, Appl
29	90	4.5	1205	9 US-10-173-123-3	Sequence 4, Appl
30	90	4.5	1207	9 US-10-173-123-4	Sequence 4, Appl
31	89.5	4.5	544	9 US-10-174-590-274	Sequence 274, App
32	89.5	4.5	544	9 US-10-176-758-274	Sequence 274, App
33	89.5	4.5	544	9 US-10-175-737-274	Sequence 274, App
34	89.5	4.5	544	9 US-10-173-706-274	Sequence 274, App
35	89.5	4.5	544	9 US-10-175-738-274	Sequence 274, App
36	89.5	4.5	544	9 US-10-175-752-274	Sequence 274, App
37	89.5	4.5	544	9 US-10-176-482-274	Sequence 274, App
38	89.5	4.5	544	9 US-10-176-757-274	Sequence 274, App
39	89.5	4.5	544	9 US-10-176-913-274	Sequence 274, App
40	89.5	4.5	544	9 US-10-180-552-274	Sequence 274, App
41	89.5	4.5	544	9 US-10-180-557-274	Sequence 274, App
42	89.5	4.5	544	9 US-10-173-700-274	Sequence 274, App
43	89.5	4.5	544	9 US-10-174-572-274	Sequence 274, App
44	89.5	4.5	544	9 US-10-174-579-274	Sequence 274, App
45	89.5	4.5	544	9 US-10-174-582-274	Sequence 274, App

#### ALIGNMENTS

RESULT 1  
US-09-764-887-176  
Sequence 176, Application US/09764887  
Patent No. US2002042096A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA113  
CURRENT APPLICATION NUMBER: US/09/764,887  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 658  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 176  
LENGTH: 108  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (104)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (105)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-887-176  
Query Match 6.2%, Score 123.5, DB 10, Length 108;  
Best Local Similarity 71.1%; Pred. No. 0.0015;  
Matches 27; Conservative 3; Mismatches 5; Indels 3; Gaps 2;  
Cy 315 ISRKGVTKLKNNGDFIANGRPP-IYIDGRPVLCGS 351  
Db 64 LSASTGVTKLKNNGDFIANGRPPSTMDGR--VCGS 99  
RESULT 2  
US-09-976-740-8  
Sequence 8, Application US/09976740  
PUDICATION No. US20020194633A1  
GENERAL INFORMATION:  
APPLICANT: Lees, Ann M.  
APPLICANT: Lees, Robert S.

APPLICANT: Law, Simon W.  
APPLICANT: Aijona, Anibal A.  
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
TITLE OF INVENTION: ATHEROSCLEROSIS  
FILE REFERENCE: 10797-004001  
CURRENT APPLICATION NUMBER: US/09/976,740  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 09/616,289  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 08/979,608  
PRIOR FILING DATE: 1997-11-26  
PRIOR APPLICATION NUMBER: US 60/031,930  
PRIOR FILING DATE: 1996-11-27  
PRIOR APPLICATION NUMBER: US 60/048,547  
PRIOR FILING DATE: 1997-06-03  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 8  
LENGTH: 530  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-976-740-8

Query Match 5.5%; Score 110.5; DB 9; Length 530;  
Best Local Similarity 18.3%; Pred. No. 0.2; Mismatches 150; Indels 111; Gaps 16;  
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVPERGRCGSEPSSEKKVSKAPSTVPSPSPAPG-LTKRVKKS-- 50  
DB 75 NCGGDEGDAQBPAPBPAEAKSRITYVANNGEPPTPVYVYGEKPSKGPNTTEIRQSD 134  
QY 51 -----KQPLQYTKDLGRKKRPANDLLINAVLQT-NDLTSVHLGVKFGSFRFLREVOEKM 103  
DB 135 VCDRDRHRRPOEKKKAGLQKE-----ITLMQTLNTLSTPEKKLAALCKKYAEILLEHR 188  
QY 104 VALLYDPVSKLACQAMROLHPEALIAIOSKALFSA----- 140  
DB 189 NSQKMKLIQKQOSQVQDE--KDHLSGHSKAVLARSKLESICRELQRNRSLSKEGVOR 246  
QY 141 ---EEOILSKVSTQSPTLETFODLLHRRPDAFYLAATAKALQAHWQ-----LMKQY 190  
DB 247 AAEFEKREKREVSHPQVTLNDIQLQMEQHNE-----RNSKLRQENNELAERLKKLIEQYE 301  
QY 191 LLEDQTVQPLPKGDQVLYNSDAEDLIDSKLQMDDEVLEHMLVADRQKEIRQLEBE 250  
DB 302 LREEH1-----DKVFKKHDLOOQLVDAKLOAOQEMLKE-----AEERHORE----- 342  
QY 251 LHKQVLYDSITGMSPPDNDQTLAVLRGMVRYLMRSREI--TLGRATKDNQIDVDLSL 308  
DB 343 --KDFLKEAVESQRMCELMKQOETHLKOQALYTEKEFEFPQNTLSKS-----EVFTTF 395  
QY 309 EGPAMKISRKQGVILKNNGDFFIANEGRRPIYIDGRPYLCSKXWRLSNNSVVEIASLRF 368  
DB 396 KQEMEMTKK--IKKLEKETTMY-----RSRWESSKALLENVAEEXT 435  
QY 369 V 369  
DB 436 V 436

RESULT 3  
US-09-962-055-8  
Sequence 8, Application US/09962055  
Patent No. US20020052033A1  
GENERAL INFORMATION:  
APPLICANT: Lees, Ann M.  
Lees, Robert S.  
Law, Simon W.  
Aijona, Anibal A.  
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/962,055  
FILING DATE: 24-Sep-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/979,608  
FILING DATE: 26-NOV-1997  
APPLICATION NUMBER: US 60/031,930  
FILING DATE: 27-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-962-055-8

Query Match 5.5%; Score 110.5; DB 10; Length 530;  
Best Local Similarity 18.3%; Pred. No. 0.2; Mismatches 150; Indels 111; Gaps 16;  
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVPERGRCGSEPSSEKKVSKAPSTVPSPSPAPG-LTKRVKKS-- 50  
DB 75 NCGGDEGDAQBPAPBPAEAKSRITYVANNGEPPTPVYVYGEKPSKGPNTTEIRQSD 134  
QY 51 -----KQPLQYTKDLGRKKRPANDLLINAVLQT-NDLTSVHLGVKFGSFRFLREVOEKM 103  
DB 135 VCDRDRHRRPOEKKKAGLQKE-----ITLMQTLNTLSTPEKKLAALCKKYAEILLEHR 188  
QY 104 VALLYDPVSKLACQAMROLHPEALIAIOSKALFSA----- 140  
DB 189 NSQKMKLIQKQOSQVQDE--KDHLSGHSKAVLARSKLESICRELQRNRSLSKEGVOR 246  
QY 141 ---EEOILSKVSTQSPTLETFODLLHRRPDAFYLAATAKALQAHWQ-----LMKQY 190  
DB 247 AAEFEKREKREVSHPQVTLNDIQLQMEQHNE-----RNSKLRQENNELAERLKKLIEQYE 301  
QY 191 LLEDQTVQPLPKGDQVLYNSDAEDLIDSKLQMDDEVLEHMLVADRQKEIRQLEBE 250  
DB 302 LREEH1-----DKVFKKHDLOOQLVDAKLOAOQEMLKE-----AEERHORE----- 342  
QY 251 LHKQVLYDSITGMSPPDNDQTLAVLRGMVRYLMRSREI--TLGRATKDNQIDVDLSL 308  
DB 343 --KDFLKEAVESQRMCELMKQOETHLKOQALYTEKEFEFPQNTLSKS-----EVFTTF 395  
QY 309 EGPAMKISRKQGVILKNNGDFFIANEGRRPIYIDGRPYLCSKXWRLSNNSVVEIASLRF 368  
DB 396 KQEMEMTKK--IKKLEKETTMY-----RSRWESSKALLENVAEEXT 435  
QY 369 V 369  
DB 436 V 436

```
RESULT 4
US-10-023-529-8
; Sequence 8, Application US/10023529
; Patent No. US2002012938A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-8

Query Match          5.5%; Score 110.5; DB 12; Length 530;
Best Local Similarity 18.3%; Pred. No. 0.2;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVPRGCSGSESSSEKKVSKAPSTVPSPSPAPG-LTKRVKKS-- 50
DB 75 NQGGGEGGAQGEPAEPDAEKSRITYVARNGEPEPTPVYGEKPSKGDPTTEIRQSD 134
QY 51 -----KQPLQVTKDLGRMKRPANDLLINAVLQT-NDLTSVHLGVKFSCTFLREVOERW 103
DB 135 VGDRDHRPRQEKKAQKAGKGE-----ITLLMQTLNTLSTPEKLAALCKYAELEBHR 188
QY 104 YALLYDPVISTLACQAMQOLHPEALIAIOSKALFSKA----- 140
DB 189 NSQOMKTLQKKQSOQVGE--KDHLRGSHSKAVLARSKLESCELOHRNRSLSKEGVQR 246
QY 141 ---EEQLLSKVGSTSQPTLETFODLHHRPDFAFYLAFTAKALQAHQ-----LMKQYY 190
DB 247 AREBEKKEKVEYSHFOVTLNDIQLOMEQHNE-----RNSKLRQEMMELAEKLLKLEIGEY 301
QY 191 LLEDVQPLPRGQOVLFNSPAEDLIDSKLDMRDEVLEHLMVADRQRKEIRQLEGE 250
DB 302 LREEHI-----DKVFGHKDLQOOQVDAKLOQAQEMLKE-----AEERHORE----- 342
QY 251 LHKQVULVDSITGMSSPFDNQTLAVLRGWRVYLMRSREI--TLGRATKXNQIDVDLSL 308
DB 343 --KQPLKEAVESQRMCMELMKQOETHLKQOLALYTEKEFEFQNTLSKSS-----EVFTTF 395
QY 309 EGPAMKISKQGVILKXNGDFPIANEGRRPIYIDGRPVLCSSKRWLSNNSVVEIASLRF 368
DB 396 KQEMEKMTKK--IKKLEKETMY-----RSRWESSNKALLLEVAEKEKT 435
QY 369 V 369
DB 436 V 436

RESULT 5
US-10-023-523-8
```

```
; Sequence 8, Application US/10023523
; Patent No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-8

Query Match          5.5%; Score 110.5; DB 12; Length 530;
Best Local Similarity 18.3%; Pred. No. 0.2;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVPRGCSGSESSSEKKVSKAPSTVPSPSPAPG-LTKRVKKS-- 50
DB 75 NQGGGEGGAQGEPAEPDAEKSRITYVARNGEPEPTPVYGEKPSKGDPTTEIRQSD 134
QY 51 -----KQPLQVTKDLGRMKRPANDLLINAVLQT-NDLTSVHLGVKFSCTFLREVOERW 103
DB 135 VGDRDHRPRQEKKAQKAGKGE-----ITLLMQTLNTLSTPEKLAALCKYAELEBHR 188
QY 104 YALLYDPVISTLACQAMQOLHPEALIAIOSKALFSKA----- 140
DB 189 NSQOMKTLQKKQSOQVGE--KDHLRGSHSKAVLARSKLESCELOHRNRSLSKEGVQR 246
QY 141 ---EEQLLSKVGSTSQPTLETFODLHHRPDFAFYLAFTAKALQAHQ-----LMKQYY 190
DB 247 AREBEKKEKVEYSHFOVTLNDIQLOMEQHNE-----RNSKLRQEMMELAEKLLKLEIGEY 301
QY 191 LLEDVQPLPRGQOVLFNSPAEDLIDSKLDMRDEVLEHLMVADRQRKEIRQLEGE 250
DB 302 LREEHI-----DKVFGHKDLQOOQVDAKLOQAQEMLKE-----AEERHORE----- 342
QY 251 LHKQVULVDSITGMSSPFDNQTLAVLRGWRVYLMRSREI--TLGRATKXNQIDVDLSL 308
DB 343 --KQPLKEAVESQRMCMELMKQOETHLKQOLALYTEKEFEFQNTLSKSS-----EVFTTF 395
QY 309 EGPAMKISKQGVILKXNGDFPIANEGRRPIYIDGRPVLCSSKRWLSNNSVVEIASLRF 368
DB 396 KQEMEKMTKK--IKKLEKETMY-----RSRWESSNKALLLEVAEKEKT 435
QY 369 V 369
DB 436 V 436

RESULT 6
US-09-976-740-44
; Sequence 44, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-744-125a-2
Query Match          5.5%; Score 110.5; DB 9; Length 546;
Best Local Similarity 18.3%; Pred. No. 0.21;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVEBERGCSGSPSSSEKKVSKAPSTVPSPSPAPG-LTKRVKYS-- 50
DB 91 NCGGPGEDGAQGEPAPEDEAKSRITYVANNGEPPTPVVNGEKEPSKGPNTETIRQSD 150
QY 51 -----KQPLQVTKDLGKWKRPANDLLINAVLOT-NDLTSVHLGVKFSRFTLREVQERW 103
DB 151 VQDRDHRPRQEKKKKGLGKE-----ITLMQNTLNTLSTPEKLAALCKYAEILLEHR 204
QY 104 VALLYDPIVSKLACQAMQLHPEALAIQSKALFSKA----- 140
DB 205 NSQKQMKLQKKQSQSLVGE--KDHLEGSHSKAVLARSKLESICRELQHNRSLEKEGVOR 262
QY 141 ---EQLSKVGSISQPTLETFODLLHRHPDAFYLAARTAKALOAHQ-----LMQY 190
DB 263 ABEERKEKREVTSHFOVLNDIQLQMEQHNE-----RNSKLRQENMELAERLKGLIEOYE 317
QY 191 LLEDQTVQPLPGDQVLPNSAEDLIDSKLKDMDDEVLEHLMVADRQKEIRQEOE 250
DB 318 LREHHI-----DKVFYHKDLOQQLVDAKLQQAQEMLKE-----AEERHORE----- 358
QY 251 LHKQVLDVDSITGMSSPDFDNTLAVLRGRMVYLMRSREI--TLGRATKNOQIDVDLSL 308
DB 359 --KDFLKEAVESQRMCELMKQOETHLKOQALALYTEKEFEEFQNTLSKS-----EVFTTF 411
QY 309 EGPAMKISRKQGVILKNNGDFFIANGRRPIYIDGRPLVCSKRLSNSVVEIASLRF 368
DB 412 KQEMEKMTKK--IKLEKETTWY-----RSRWSSNKALLMAAEKRT 451
QY 369 V 369
DB 452 V 452

RESULT 7
US-10-023-529-44
; Sequence 44, Application US/10023529
; Patent No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
```

```

; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-529-44
Query Match          5.5%; Score 110.5; DB 12; Length 546;
Best Local Similarity 18.3%; Pred. No. 0.21;
Matches 77; Conservative 83; Mismatches 150; Indels 111; Gaps 16;

QY 1 NSARGVEBERGCSGSPSSSEKKVSKAPSTVPSPSPAPG-LTKRVKYS-- 50
DB 91 NCGGPGEDGAQGEPAPEDEAKSRITYVANNGEPPTPVVNGEKEPSKGPNTETIRQSD 150
QY 51 -----KQPLQVTKDLGKWKRPANDLLINAVLOT-NDLTSVHLGVKFSRFTLREVQERW 103
DB 151 VQDRDHRPRQEKKKKGLGKE-----ITLMQNTLNTLSTPEKLAALCKYAEILLEHR 204
QY 104 VALLYDPIVSKLACQAMQLHPEALAIQSKALFSKA----- 140
DB 205 NSQKQMKLQKKQSQSLVGE--KDHLEGSHSKAVLARSKLESICRELQHNRSLEKEGVOR 262
QY 141 ---EQLSKVGSISQPTLETFODLLHRHPDAFYLAARTAKALOAHQ-----LMQY 190
DB 263 ABEERKEKREVTSHFOVLNDIQLQMEQHNE-----RNSKLRQENMELAERLKGLIEOYE 317
QY 191 LLEDQTVQPLPGDQVLPNSAEDLIDSKLKDMDDEVLEHLMVADRQKEIRQEOE 250
DB 318 LREHHI-----DKVFYHKDLOQQLVDAKLQQAQEMLKE-----AEERHORE----- 358
QY 251 LHKQVLDVDSITGMSSPDFDNTLAVLRGRMVYLMRSREI--TLGRATKNOQIDVDLSL 308
DB 359 --KDFLKEAVESQRMCELMKQOETHLKOQALALYTEKEFEEFQNTLSKS-----EVFTTF 411
QY 309 EGPAMKISRKQGVILKNNGDFFIANGRRPIYIDGRPLVCSKRLSNSVVEIASLRF 368
DB 412 KQEMEKMTKK--IKLEKETTWY-----RSRWSSNKALLMAAEKRT 451
QY 369 V 369
DB 452 V 452

RESULT 8
US-10-023-523-44
; Sequence 44, Application US/10023523
; Patent No. US2002012485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
```



;; CURRENT APPLICATION NUMBER: US/10/023,523  
;; CURRENT FILING DATE: 2001-12-17  
;; PRIOR APPLICATION NUMBER: US/09/616,289  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 09/517,849  
;; PRIOR FILING DATE: 2000-03-02  
;; PRIOR APPLICATION NUMBER: US 08/979,608  
;; PRIOR FILING DATE: 1997-11-26  
;; PRIOR APPLICATION NUMBER: US 60/031,930  
;; PRIOR FILING DATE: 1996-11-27  
;; PRIOR APPLICATION NUMBER: US 60/048,547  
;; PRIOR FILING DATE: 1997-06-03  
;; NUMBER OF SEQ ID NOS: 53  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 44  
;; LENGTH: 546  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-023-523-44

Query Match 5.5%; Score 110.5; DB 12; Length 546;  
Best Local Similarity 18.3%; Pred. No. 0.21; Indels 111; Gaps 16;  
Matches 77; Conservative 83; Mismatches 150;

QY 1 NSARGVPEPGCGSGEPSSSEKKVSKAPSTVPSPSPAPG-LTKVYKKS-- 50  
DB 91 NQGRPEAGGAEPAEPDAEKSRITYVANGEPPTPVNGEKEKSKDPTTEIRQSD 150  
QY 51 -----KQPLQVTKDGLRWKRPANDLLINAVLQT-NDLTSVHLGVKFSCTFLREVOERW 103  
DB 151 VGDRDHRPQEKKAQGLGKE-----ITLMQTLNTLSTPEEKALCKKYAELLEHR 204  
QY 104 YALYDPVTSKLAQAMQHPAIAAQSKALFSA----- 140  
DB 205 NSQKMKLQKQSQQLVQE--KDHLRGHSKAVLARSKLSLCRELQHNHNSLKEGVOR 262  
QY 141 ---EEQLSKTGSTSOPLETFTQDILHRPDAFYLAFTAKALQAHWQ-----LMKOYU 190  
DB 263 ARREEKREKVTSHQVTLNDIQLOMEQHNE-----RNSKLQEMELAEKLIQYELREHI 317  
QY 191 LLEDOVQPLPKGDQVLNPSDAEDLIDSKLQDMDEVLEHMLVADRROREIRQLEOE 250  
DB 318 LREHI-----DKVFYGHKDLQQLVDAKLQQAQEMLKE-----AEERHQRE 358  
QY 251 LHKQVLDVSTGSSPPFDNQTLAVLRGMRYLMRSREI--TLGRATKDNQIDVDLSL 308  
DB 359 --KQFLKEAVESQRMCELMQOETHLQKQALALYTEKEFEFQNTLSKSS-----EVFTTF 411  
QY 309 EGPAWKISRKQGVITLKNNGDPFIANEGRRPIYIDGRPVLCGSKWRLSNNSVVEIASLRF 368  
DB 412 KQNEKMTKK--IKLEKETMY-----RSRWESSNKALLEMAEEXT 451  
QY 369 V 369  
DB 452 V 452

RESULT 9  
US-09-976-740-5  
;; Sequence 5, Application US/09976740  
;; Publication No. US20020194633A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lees, Ann M.  
;; APPLICANT: Lees, Robert S.  
;; APPLICANT: Law, Simon W.  
;; APPLICANT: Arjona, Anibal A.  
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
;; TITLE OF INVENTION: ATHEROSCLEROSIS  
;; FILE REFERENCE: 10797-004001  
;; CURRENT APPLICATION NUMBER: US/09/976,740  
;; CURRENT FILING DATE: 2001-10-12  
;; PRIOR APPLICATION NUMBER: 09/616,289

;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 08/979,608  
;; PRIOR FILING DATE: 1997-11-26  
;; PRIOR APPLICATION NUMBER: US 60/031,930  
;; PRIOR FILING DATE: 1996-11-27  
;; PRIOR APPLICATION NUMBER: US 60/048,547  
;; PRIOR FILING DATE: 1997-06-03  
;; NUMBER OF SEQ ID NOS: 53  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 5  
;; LENGTH: 557  
;; TYPE: PRT  
;; ORGANISM: Oryctolagus cuniculus  
US-09-976-740-5

Query Match 5.4%; Score 107; DB 9; Length 557;  
Best Local Similarity 18.9%; Pred. No. 0.44;  
Matches 76; Conservative 76; Mismatches 137; Indels 114; Gaps 17;

QY 16 EPSSEKKVSKA-----PSTPV-----PSPAPAGILTKRVKKS-----KQPLQVT 57  
DB 105 EPDAEKSRAYVANGEPPTPVNGEKEKSKAPG--TEIRTSDEVGDHRRPQEK 163  
QY 58 KDLGRWKPANDLLINAVLQT-NDLTSVHLGVKFSCTFLREVOERYALLYDPVSKLA 116  
DB 164 KAKGLGKE-----ITLMQTLNTLSTPEEKALCKKYAELLEHRNSQKMKLQKQ 217  
QY 117 COMRQLHPEAIAAQSKALFSA-----BEQLSKYGS 150  
DB 218 SOLVQE--KHLRREHSKAIARSKLSLCRELQHNHNSLKEGVORAEERKEKVT 275  
QY 151 TSQPLTFQDILHRPDAFYLAFTAKALQAHWQ-----LMKOYLLPDQVQPLPKG 203  
DB 276 HFQMTLNDIQLOMEQHNE-----RNSKLQEMELAEKLIQYELREHI 323  
QY 204 DQVNLFSDAEDLIDSKLQDMDEVLEHMLVADRROREIROELHKKQVLDVSTG 263  
DB 324 DKVFYGHKDLQQLVDAKLQQAQEMLKE-----AEERHQRE-----KQFLKEAVES 369  
QY 264 MSSPFDNQTLAVLRGMRYLMRSREI--TLGRATKDNQIDVDLSLEGPWKISRKGV 321  
DB 370 QRMCELMQOETHLQKQALALYTEKEFEFQNTLSKSS-----EVFTTFQKQEMKTKK--I 422  
QY 322 IKLKNNGDPFIANEGRRPIYIDGRPVLCGSKWRLSNNSVVEIA 364  
DB 423 KLEKETMY-----RSRWESSNKALLEMA 447

RESULT 10  
US-09-962-055-5  
;; Sequence 5, Application US/09962055  
;; Patent No. US20020052033A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lees, Ann M.  
;; APPLICANT: Lees, Robert S.  
;; APPLICANT: Law, Simon W.  
;; APPLICANT: Arjona, Anibal A.  
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
;; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
;; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS  
;; NUMBER OF SEQUENCES: 42  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson P.C.  
;; STREET: 225 Franklin Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02110-2804  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/09/962,055
3 FILING DATE: 24-Sep-2001
4
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: 08/979,608
7 FILING DATE: 26-NOV-1997
8 APPLICATION NUMBER: US 60/031,930
9 FILING DATE: 27-NOV-1996
10
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Myers, Louis
13 REGISTRATION NUMBER: 35,965
14 REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 617/542-5070
17 TELEFAX: 617/542-8906
18
19 INFORMATION FOR SEQ ID NO: 5:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 557 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25 FRAGMENT TYPE: internal
26 SEQUENCE DESCRIPTION: SEQ ID NO: 5:
27
28 US-09-962-055-5
29
30 Query Match 5.4%; Score 107; DB 10; Length 557;
31 Best Local Similarity 18.9%; Pred. No. 0.44;
32 Matches 76; Conservative 76; Mismatches 137; Indels 114; Gaps 17
33
34
35 16 EPPSEKKKKVSKA-----PSTPV-----PPSPAPAPGLTKRVKKS-----KQPLQVT 57
36 105 EPEDEKSRAYARRNGEPPEPTPVNGEKETSKAEPG-TETRTSDVGDROHRRPQEK 163
37 58 KDLGRWKPANDLLINAVQT-NDLTSVHGVKFSCTFLREVERWVALLYDPVSKLA 116
38 164 KAKGIGKE-----ITLMQTLNTLSTPEKTLALCKYVAELLEHRNSQKQKLLQKQ 217
39 117 CQAMQLHDEAIAIOSKALFSKA-----EEQLLSKVG 150
40 218 SOLVOE--KDLRGESKAIILARSKLESCLREIQRHNSLKEGVORAREEKEKREKVS 275
41 151 TSQPLTFEQDILHHHPDAFYLAFTAKALQAMQ-----LMKQYLLDEDQVQLPFG 203
42 216 HFQMTLNDIQLOMEQNE-----RNSKLRQENNELAERLKLIEYELREHI----- 323
43 204 DQVLNPSDAEDLIDSKLQMDREDEVLEHMAVDRROKKEIRQLBQLHKQVLYDSIG 263
44 324 DKVPFKHKLQQLVADKALQAOEMLKE-----AEERHQE-----KDLFLKAVS 369
45 264 MSSPFDNQTLAVLRGRWVYLMRSHEI--TLGRATKQNOIDVLSLEGPAMKISKQGV 321
46 370 QMCELMKQOETHLKQQLALYTEKFEFPONTLSKSS-----EVFTTFQOEMEKTKK-I 422
47
48 322 IYLNKNGDFFIANEGRRPIYIDGRPVLGSSKRLSNISVVELA 364
49 423 KLEKETTVY-----RSRWESSKALLEMA 447
50
51 RESULT 11
52 US-10-023-529-5
53 Sequence 5, Application US/10023529
54 Patent No. US20020129388A1
55 GENERAL INFORMATION:
56 APPLICANT: Lees, Ann M.
57 APPLICANT: Lees, Robert S.
58 APPLICANT: Lees, Simon W.
59 APPLICANT: Arjona, Anibal A.
60 TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
61 TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
62 TITLE OF INVENTION: ATHEROSCLEROSIS
63 FILE REFERENCE: 10797-004001
64 CURRENT APPLICATION NUMBER: US/10/023,529
65 CURRENT FILING DATE: 2001-12-17

```

```

      / PRIOR APPLICATION NUMBER: 09/616,289
      / PRIOR FILING DATE: 2000-07-14
      / PRIOR APPLICATION NUMBER: US 09/517,849
      / PRIOR FILING DATE: 2000-03-02
      / PRIOR APPLICATION NUMBER: US 08/979,608
      / PRIOR FILING DATE: 1997-11-26
      / PRIOR APPLICATION NUMBER: US 60/031,930
      / PRIOR FILING DATE: 1996-11-27
      / PRIOR APPLICATION NUMBER: US 60/048,547
      / PRIOR FILING DATE: 1997-06-03
      / NUMBER OF SEQ ID NOS: 53
      / SOFTWARE: FastSeq for Windows Version 4.0
      / SEQ ID NO: 5
      / LENGTH: 557
      / TYPE: PRT
      / ORGANISM: Oryctolagus cuniculus
US-10-023-529-5

Query Match          5.4%; Score 107; DB 12; Length 557;
Best Local Similarity 18.9%; Pred. No. 0.44; Indels 114; Gaps 17
Matches    76; Conservative    76; Mismatches 137;

Qy   16 EPPSEKKKVKSA-----PSTPV-----PPSPAPAGLTIRYKS-----KOPLOVT 57
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   105 EEDEDEKRAYANRNGEPEPTPVNNGEKETSKAPG--TEIIRTSDEVODNRHRRPEKK 163
Qy   58 KDLGRMKRPANDILLINAVLQT-NDITSVHLGVKFSCEFTLRVEQRMWALYDPVYSKLA 116
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   164 KAKGLGKE-----ITLLMQTLNTLSTPEBKLAALCKKYAEELLEHRNSQKOMKLQKQ 217
Qy   117 QAMQMQLHEAIALAQSKALPSKA-----EROLLSKVGS 150
     :::::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   218 SOLVGE--KDRIHGSHSKAILARSKLSLCRELQHNRSLKEEGVQRAREEEKREKETS 275
Qy   151 TSQPLETEFODILHRHPAFYLATPAKALOAWQ-----LMKOYULLEDOTVOPLPKG 203
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   276 HQMTLNIDIQLOMEQNE-----RSKIKRQEMELAEARKKUIEQELYELREHI----- 323
Qy   204 DQVLNFPDAEDLIDDSKLDKMDREDEVLEHLMVADRROKREIRQLHQELHKWOVLVDSTIG 263
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   324 DKVFVGHKDIQQQLVDAKIQAOQEMLKE-----AEEBHORE-----KOFLLKEAVES 369
Qy   264 MSSPFDNQTLAVLRGRWRYILMSRELI--TLGRATXNQNDIVDSLSEGPARIKRKGCV 321
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   370 GRMCIMKQOETHLKQOALALYTEKEEFQNTLSKSS----EVFTTFQOEEMKTKK--I 422
Qy   322 ICLKNNGDPEFIANGRRPIYIDGRPVLGSGKRWLISNNSVGEIA 364
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   423 KLLEKETMY-----RSRWESSNKALLERA 447

RESULT 12
US-10-023-523-5
/ Sequence 5, Application US/10023523
/ Patent No. US20020152485A1
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ APPLICANT: Lees, Robert S.
/ APPLICANT: Law, Simon W.
/ APPLICANT: Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
/ TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
/ FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/10/023,523
/ CURRENT FILING DATE: 2001-12-17
/ PRIOR APPLICATION NUMBER: US/09/616,289
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: US 08/979,608
/ PRIOR FILING DATE: 1997-11-26
/ PRIOR APPLICATION NUMBER: US 60/031,930
```

PRIOR FILING DATE: 1996-11-27  
PRIOR APPLICATION NUMBER: US 60/048,547  
PRIOR FILING DATE: 1997-06-03  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 557  
TYPE: PRT  
ORGANISM: Oryctolagus cuniculus  
US-10-023-523-5

Query Match 5.4%; Score 107; DB 12; Length 557;  
Best Local Similarity 18.9%; Pred. No. 0.44;  
Matches 76; Conservative 76; Mismatches 137; Indels 114; Gaps 17;

QY 16 EBSSEKKKVSKA-----PSTPV-----PPSPAPAGLTKRVKS-----KQPLQVT 57  
DB 105 EPDEKSRAYAVRANGEPPTPVVNGEKETSKAPG--TEETRTSDEVGDRDRRPPQEK 163  
QY 58 KDLGRWKANDLLILNAVLT-NDLTSVHLGVFSCFPTLREVSQRYALLVDPVISKLA 116  
DB 164 KAKGIGKE-----ITLMQTLNTLSTPEBKLAALCKKAYELIEHRNSQKQMKLQKKQ 217  
QY 117 QAMQQLPEALIAIQSKALFSKA-----EEOQLSKYGS 150  
DB 218 SQLVGE--KDLRGESHKAILARSKLESICRELQRHNSLKEGVQARAEESEKKKEVTS 275  
QY 151 TSQPLTFPQDLHHPDAFYIARTAKALQAHWQ-----LMKQYLLDQTVPLRK 203  
DB 276 HFQMTLNDIQLQMEQHNE-----RNSKLRQENMELEARKKLEIQEYELREHEI----- 323  
QY 204 DQVLFSDAEDLIDSKLKDMDDEVLEHLMVADBRQKEIRQLQELHMKQVLVDSITG 263  
DB 324 DKVFYKHLQOOLVDKQLQAOEMLKE-----AERHORE-----KDFLKEAVES 369  
QY 264 MSSPFDNQTALVLRGRVYLMRSREI--TVGRATKQNDVDLSLSPAWKISKQGV 321  
DB 370 QRMGLMKQOETHLKOQLALYTEKEEFQNTLSKSS-----EVFTTFQOEMKMTKK--I 422  
QY 322 IYKLNKNGDFIANEGRRPIYIDGRVYLCGSKKRLSNNSVVEIA 364  
DB 423 KLEKETMY-----RSRWESSNKALLEMA 447

RESULT 13  
US-09-833-790-435  
Sequence 435, Application US/09833790  
Patent No. US20020068288A1  
GENERAL INFORMATION:  
APPLICANT: Lodes, Michael J.  
APPLICANT: Wang, Tongtong  
APPLICANT: Secrist, Heather  
APPLICANT: Mohamath, Radooh Y.  
APPLICANT: Indrias, Carol Y.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.512  
CURRENT APPLICATION NUMBER: US/09/833,790  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 435  
LENGTH: 519  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-790-435

Query Match 5.3%; Score 106; DB 10; Length 519;  
Best Local Similarity 21.8%; Pred. No. 0.49;  
Matches 87; Conservative 57; Mismatches 144; Indels 112; Gaps 20;  
QY 62 RWKPANDLLILNAVLT-----QTNDLTSV--HIGVKSGRFTLREVOERWYALLVDPVISK 114

DB 8 RTMPAVDQLLEBALQDSPQTRSLSVFEEDAGTLTDYTNQLLOAMQRYGAONE----- 62  
QY 115 LACQMRQLHDEALIAIQSKALFSKAEQQLLSKXGTSQPLTFPQDLHHPDAFYIAR 174  
DB 63 -MCLATQQLSKQALAYEKONFALGKGDEEVS-----TLH-----YFSK 100  
QY 175 TAKALQ-AHWQLMQRYLLLEDQTVQPLPKGDQVLFNSDAEDLIDSKLKDMDDEVLT-BHE 232  
DB 101 VVDELINLHTELARQ---LADTWLP-----IIQFRE-KOLTEVSTLAKDLFGLASNEHD 150  
QY 233 LM-----VADRRQKREIRQLEQ-----ELHKQVLVDSIT 262  
DB 151 LSMAYSRLLPKKENEKVTEGVKEVAARRKQHLSSIQYCALNALQYRKQAMMBPMI 210  
QY 263 GMS--SPDFDNTALVLRGRVRL-----MRSREITLG-----RATKQNDVDLSL 308  
DB 211 GFHQQINFQKGAEMFSKRMDSFLSVADWQSTQVELAEAEKMRVSOQELLSDVESV 270  
QY 309 EGP-----AMKISR-----KQGVIKLKNNGDFIANEGRRPIYIDGRVYLC-----GS 351  
DB 271 YTPSDVAAPQINRNLQKAGYINLRKNTGLVTTTWERLYFTTQGNLMQPRGAVAGL 330  
QY 352 KWRLSNNSV--VEIASLRFVLI---NODLIALIAEBAK 386  
DB 331 IQDLNCSVMAVDCEDRRYCFQITTPNGKSGIILQAESRK 370

RESULT 14  
US-09-864-761-38419  
Sequence 38419, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687

```

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38419
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049872.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.72
; OTHER INFORMATION: EST HUMAN HIT: AU138211.1, EVALU2 2.00e-83
; OTHER INFORMATION: SWISSPROT HIT: Q10411, EVALU2 3.00e-18
US-09-864-761-38419

```

```

Query Match      5.1%; Score 102.5; DB 10; Length 751;
Beet Local Similarity 20.4%; Pred. No. 1.7;
Matches 84; Conservative 70; Mismatches 128; Indels 129; Gaps 19;

```

```

QY 18 SSSSEKKYKSKAPSTVPSPAPAPGTLTKR-----VKKSKQPLQYTDL-LGRMKP 65
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 237 NSEKHKDLSKAEKTERL-----VKGIKERLEIKLNEKNISLTKQIDQLSKDEVGK--- 287
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 ANDLLINAVLQTNDLTSVHLGVKSCFPTLREVV-----QBRVYALLYVPVSKLACQAM 120
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 288 -----LTQIQKQDEIGALHARISSTHTQVUVYLLQOQLQYANERKKAVALNEKTR 341
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 ROLHREA-----IAIOLSKALFSKAEK--OLSKVSGTSQPTL-ETFODL-LHRHPD 168
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 342 ESHHKTEYHKMDIVAAKEAALIKLODENKKLSTPESGQDMFRETIOVLSRIIREKD 401
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 AFY--LAPAKALQAHWOLMKOYLLLEDQTVQPLPKGDV--LNSDAEDLIDD-SKLKD 223
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 402 IEIDLQSCQTLA-----VLQTSSTGNAGVNSQPFELLQERDKLQ 447
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 M--RDEVLEHEIMVADRQREIRLOEELHKW--QVLVDSITGMSSPDFDQTLAVLKG 279
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 448 QVKKEWKOQVVTYQNMQHSQAQLEELHQLOAVLVDS-----DNNSKL----- 494
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 280 RMVRYLMSREITLGRATKQNDIVDLSLEGPAWKISRQGVYKLNQNGDFPANEGRBP 339
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 495 -----QVDYTGILQ-----SYQNETKCLKN----- 514
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 340 IYIDGRPVLCGSKWRLSNNSVVEIASLRFVFLINDLILAPAAKTIPTQ 390
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 515 -----FGQELAQVQHSIGQLCNTKDLILGKDLIIPQSSASLTPQ 556
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 15
US-09-815-242-11259
; Sequence 11259, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyckind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard

```

```

; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11259
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11259

```

```

Query Match      5.0%; Score 100.5; DB 10; Length 475;
Beet Local Similarity 19.1%; Pred. No. 1.3;
Matches 84; Conservative 62; Mismatches 136; Indels 157; Gaps 18;

```

```

QY 14 GSEBSSSEKKYKSKAPSTVPSPAPAPGTLTKRVKSKQPLQYTDLG----- 61
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 7 GKLSTANTEMKAKAPKT-----AYVNDQGAESFRWQGGCSA 44
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 --RMKRPANDLLINAVLQTNDLTSVHLGVKSCFPTLREVOERWYALLYDPVYSKLACQA 119
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 45 CKAANTTIEVRLISLAKSKNDPFGSAGETQAKIOTLSEI----- 84
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 MCOLHPEAIAIOLSKALFSKAEOLSKVSGTSQPTLETFODLHHRPDAYLARTAKAL 179
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 -----SLQETPRRFSGSELDRLVGGIVGSAI---LIGHPCA---GKSTILL 128
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 QAHNOLMKQ---YLLLEQTVQPLPKGDVYLFNSAEDLIDSKIKMRDEVLEHEIMWA 236
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 129 QVMGLAKMTALVYTGEBSLQOVAMRASRLG-----LPSDQKMLSETSVBOICNLA 181
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 DRQKREIROLEOEELHKWQVLVDSITGMSSPDFDQTLAVLGR-----MVRYLMSREI 291
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 182 D-----QLKQPI-----IYVDSIQVHNLADIQSSPQSVAGVRECASEFLTRY AKTRQV 228
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 292 TL--GRATKD-----NQIDVDLSLEGPA--WKISR----- 317
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 229 AIIWGHVTKDGTLAGPKVLEHAIDCSLLLEGADSRYRTLRSHKRPAGAVNELGVFGWT 288
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 318 KQGVYKLNQNGDFPANEGRPIYIDGRPVLCGSK-----WKLSNNSVVEIANSI 366
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 289 EQLGAEVKNPSAIFDSR-----GDEITSGSSVMVWLBGTRPLVLEIQALVDHSMILA 339
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 367 ---RFVFLINDLILALIRA 382
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 340 NRRVAVGLEQNRALDLIA 358
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Search completed: March 27, 2003, 05:29:56
Job time : 34 secs

```

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2003, 05:15:30 ; Search time 41 Seconds  
(without alignments)  
914.449 Million cell updates/sec

Title: US-09-744-125A-2

Perfect score: 1996

Sequence: 1 NSARGVGPGRGSGSEPPSSS.....LINODLIATRAAKITPQ 390

Scoring table: BIOSUM62

Gap: 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	231.5	11.6	408	2 T23127	hypothetical prote
2	103	10.3	735	2 T47594	hypothetical prote
3	145	7.1	551	2 D96632	hypothetical prote
4	116.5	5.8	642	2 T39607	fork head protein
5	116	5.8	676	2 AE2417	hypothetical prote
6	116	5.8	2346	2 T33829	Tpr homolog - frui
7	115	5.8	699	1 S38982	kinesin-related pr
8	111	5.6	804	2 A96494	protein F7F2.16 [i
9	106.5	5.3	1440	2 G89113	protein ZK742.1 [i
10	106.5	5.3	4574	2 G02520	plectin - human
11	106.5	5.3	3684	2 A59404	plectin [imported]
12	106	5.3	3685	1 A27605	dystrophin, muscle
13	104.5	5.2	4588	1 T28667	dystrophin beta, heavy
14	103.5	5.2	558	2 G96522	FLN1A17.16 [impor
15	103.5	5.2	1825	2 T42725	actin binding prot
16	103.5	5.2	1885	2 T30847	actin binding prot
17	103.5	5.2	2033	2 T30849	actin binding prot
18	103	5.2	810	2 B70154	DNA gyrase chain A
19	102	5.1	411	2 T55223	hypothetical prote
20	102	5.1	968	2 T51523	clpB heat shock pr
21	102	5.1	3678	2 S28916	dystrophin - mouse
22	101.5	5.1	1109	1 B45345	HIV-1 retropepsin
23	101	5.1	943	2 S68824	rngB protein, cyto
24	100.5	5.0	475	1 T64131	DNA repair protein
25	100.5	5.0	853	2 T10591	hypothetical prote
26	100	5.0	407	2 T08496	probable transcript
27	100	5.0	572	2 H86257	protein F501.2 [i
28	100	5.0	859	2 T01461	hypothetical prote
29	100	5.0	971	2 T10678	hypothetical prote

30	100	5.0	2223	2 S65074	pyrimidine synthes
31	100	5.0	15281	2 S41309	cyclosporin synthe
32	99.5	5.0	417	2 E64200	serine-tRNA ligase
33	99.5	5.0	496	2 A47199	1-aminocyclopropan
34	99.5	5.0	911	2 F96771	heat shock protein
35	99.5	5.0	1490	2 S32373	DNA-binding protei
36	99.5	5.0	2068	2 A47371	transcription init
37	99	5.0	2244	2 T1616	carbamoyl-phosphat
38	98.5	4.9	1205	2 D83862	hypothetical prote
39	98	4.9	743	2 H96493	protein F7F2.18 [i
40	98	4.9	1278	2 T27925	hypothetical prote
41	97.5	4.9	677	2 H64574	DNA topoisomerase
42	97.5	4.9	835	2 AD2441	endopeptidase Clp
43	97.5	4.9	958	2 T10679	hypothetical prote
44	97.5	4.9	1048	1 BVESC3	exonuclease (EC 3.
45	97.5	4.9	4644	1 A38905	dystrophin heavy chain

## ALIGNMENTS

## RESULT 1

T23127 hypothetical protein H28016.2 - Caenorhabditis elegans

C1:Species: Caenorhabditis elegans

C1:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C1:Accession: T23127

R:White, S. submitted to the EMBL Data Library, June 1998

A:Accession: T23127

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-408 <WIL>

A:Cross-references: EMBL:AL023815; PIRN:CAA19428.1; GSPDB:GM00019; CBSP:H28016.2

A:Experimental source: clone H28016

A:Gene: CBSP:H28016.2

A:Map position: 1

A:Introns: 39/2; 62/3; 89/3; 254/3; 268/2; 321/2; 402/2

Query Match 11.6%; Score 231.5; DB 2; Length 408;

Best Local Similarity 21.6%; Pred. No. 7.7e-09;

Matches 87; Conservative 69; Mismatches 134; Indels 113; Gaps 12;

QY	9	PERCGSPSSSEKKVSKAPSTP-----VPSPAPAPGLTK--YKSKQPLQVTKD 59
DB	44	PNELSGFKKOKIEENSEIHAETESVEAOWMKKPEPELKRKILIHNEPSITETNP 103
QY	60	LGRMKPANDLLINAVLOTNDLTSVHLGVKFCRTLAEGVERWYALLYPVISLACQA 119
DB	104	KRIWVADDLALTLTVAHVOCIRFLHNSLTPSRKTFSDVEERYQLMDEIISKAKVR 163
QY	120	MROLHPAIIAIOAKALFSKAE----- 142
DB	164	LDAMHRLKAQIEATPTTRNEERTIMELAENOLKNARKQDVENHTVLTILHKKILID 223
QY	143	-----QLSKVGSTSQPT-----LETFODLLHHPDAFYLAATKALQA 181
DB	224	GNRVYSKFSKAKICLAELIVQIGSTQKAKSGLFRQKYQ--TFKEKTQSFHKSRTQVLS 282
QY	182	HMQLMKQVYLLLEDQVQLPFGDQVLFNS--DAEDLIDSKLQKMDVLEHLMVADBRQ 240
DB	283	HYRRIKGY-----RSEGNNNNMQLLEALTLGNTWDFD--INAPLQARSRY 327
QY	241	KREIRQ--LEQELHKQVVLVDSITGMSPPDNQTLAVLRGMVYLRSEITLGRATK 298
DB	328	AAISRRPLSLGLNRFK-----TSGSVDP--NAIMINIQFLQYANTGSGVTMGRASTL 378
QY	299	DNQIDVDVLSLEGPAWKISRKQGVIKLKNQDFFLANBGRRPY 341
DB	379	NEKIDIDLSKSGPARKVL-----EGRROLF 403

## RESULT 2

hypothetical protein T12E18.40 - Arabidopsis thaliana  
T47594  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47594  
R:Blocker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes,  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: 224469  
A:Accession: T47594  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-735 <BL0>  
A:Cross-references: EMBL:AL132971  
A:Experimental source: cultivar Columbia; BAC clone T12E18  
C:Genetics:  
A:Map position: 3  
A:introns: 27/2; 56/3; 588/1; 637/3; 666/3; 709/3  
A:Note: T12E18.40

Query Match 10.3%; Score 205; DB 2; Length 735;  
Best Local Similarity 32.6%; Pred. No. 1.3e-06;  
Matches 56; Conservative 35; Mismatches 63; Indels 18; Gaps 4;

OY 208 NNSDADLIDDSKLR-DMDR---EVLHEHLMVADRQKEIRQLEQLHKQVLDVDSIT 262  
DB 563 NNSDIEAMTLDLDPDDNDFLEVSQY---SQDMKRTIIRLQAAHSYQRAISR 618  
OY 263 GMSPPDFDQTLAVLRGRWRYLMSREITLGRATKNOIDVDLSLGPAMKISRQGYI 322  
DB 619 G-----AAVLYGRSKHYIKKEPVVGRSTEDLAVDIDLGKSKSRQAI 669  
OY 323 KLNNGDFPIANEGRRPIYIDGRPVLCGSKMRLSNNSVVEIASLRFVFLNI 374  
DB 670 RLGDGSEFHIKRLGKYSISVNEKEVDPCQSLILKSDCLVEIRGMPFIETNQ 721

## RESULT 3

hypothetical protein F8A5.22 [imported] - Arabidopsis thaliana  
D96632  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: D96632  
R:Thelodis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hultzar, L.  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salber, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A66141; MUID:21016719; PMID:11130712  
C:Accession: D96632  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-551 <STO>  
A:Cross-references: GB:AB005173; NID:g2462740; PIDN:AB71959.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F8A5.22  
A:Map position: 1

Query Match 7.1%; Score 142; DB 2; Length 551;  
Best Local Similarity 23.5%; Pred. No. 0.025;  
Matches 47; Conservative 38; Mismatches 73; Indels 42; Gaps 4;

OY 211 DABDLID-DSKLR-----DMRDEVLEHLMVADRQKEIRQLEQLHKQVLDVDSITG 263  
DB 339 DEEBEIDIAMIRKTLNLPVDDSDCFNRREWNMSKPRHALLGLEO-----CTRTS 389  
OY 264 WSPDDNDQTLAVLRGRWRYLMSREITLGRATKNOIDVDLSLGPAMKISRQGYI 323

DB 390 MGRALMFHGAIALVLCPPSKHFVRKREVLIGRSSGGLANDIDLGKYNVSKISRQALVK 449  
OY 324 LKNNGDFPIANEGRRPIYIDGRPVLCGSKMRLSNNS----- 359  
DB 450 LENYGSFSLKNGKOHILVNGGKLDGRQIVTLTSLGNSINVTSLCFVQVCFIPLSTRT 509  
OY 360 --VVEIASLRFVFLINODLI 377  
DB 510 AQLIQIRGITTFVKINKAV 529

## RESULT 4

fork head protein type transcription factor - fission yeast (Schizosaccharomyces pombe)  
T39607  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39607  
R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.  
submitted to the EMBL Data Library, May 1998  
A:Reference number: 221866  
A:Accession: T39607  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-642 <LYN>  
A:Cross-references: EMBL:AL023554; PIDN:CAA19034.1; GSPDB:GN00067; SPDB:SPBC16G5.15C  
A:Experimental source: strain 972h-, cosmid c16G5  
C:Genetics:  
A:Gene: SPDB:SPBC16G5.15C  
A:Map position: 2  
A:introns: 171/2

Query Match 5.8%; Score 116.5; DB 2; Length 642;  
Best Local Similarity 24.0%; Pred. No. 1.9;  
Matches 43; Conservative 42; Mismatches 77; Indels 17; Gaps 6;

OY 210 SDAEDLIDDSKLRDMDR---EVLHEHLMVADRQKEIRQLEQLHKQVLDVDSITGM 264  
DB 8 SKSEHSDDERKQDLYKKQMDVDTRNIVLNGRLSGLATSLVPRHEMRVVDYDSNKK 67  
OY 265 SSPDFDN-QTLAVLRGRWRYLMSREITLGRAT-----KDNQIDV-DLSLGPAMK 314  
DB 68 NMRHSGEIQAYAKPAGSTWYVKKIRILLGREPANSPKKNEDLEVYDMNF-GPSKV 126  
OY 315 ISRKQCVTLK-LNNGDFPIANEGRRPIYIDGRPVLCGSKMRLSNNSVVEIASLRFVFLI 372  
DB 127 VSRKIAVEYDLDQTNCSVYGRNGIKVDGKLFNGETVKLTSGSILEVAGLQMMFVL 185

## RESULT 5

hypothetical protein alr4893 [imported] - Nostoc sp. (strain PCC 7120)  
AE2417  
C:Species: Nostoc sp.  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AE2417  
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriuch,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
C:Accession: AE2417  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-676 <KUN>  
A:Cross-references: GB:BA000019; PIDN:BA76592.1; PID:917134030; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr4893

Query Match 5.8%; Score 116; DB 2; Length 676;  
Best Local Similarity 19.8%; Pred. No. 2.2;  
Matches 69; Conservative 64; Mismatches 126; Indels 90; Gaps 14;

```

QY 5 GVEPGRCGSGSPSSSEKKYVK-----APSTP---VPPSPAPAGLTGRVYKSKQPLQV 56
D 53 GNVTPPEPEVTPPQSTKSRVIAQNSPVLPSTPRVPVPAPRATVNE-----LVV 104
QY 57 TKDGRMKRANLILLINAVLQTNLITSVHLGVKFSCTFLREVGGRWYALLVDPVSKL- 115
D 105 T-----ATDVQIVGATPELQEIIRQVVKIQTGSDITSQTLQDVAALLETGTFASAN 156
QY 116 -----ACQAMQQLHEAIAAIG---SKAL-PSKAEQQLSKVGSQPLTFQDLL 163
D 157 VNSRTTSGLVNVVQVQVIVASLQTLGAKALITYVAQPRFQSQIGKPISP-----EGL- 210
QY 164 HRHPDAFLATATKALQAMQQLMQLKQYLLLEDQTVPLPKGQDVLFSPDAEDLIDSKLKD 223
D 211 -----KQAVQAVNQWADNGYNLARVLSIFPNQG--ILNINVAEGVLSIKRFRF 258
QY 224 MRDEVLEHLMVADRROKREIRQLEQELHKQVAVDS-----ITGSSSPDPNQTAVLRG 279
D 259 VNDD-----GKTIDSNQNPVGKRTKPDPLRQQLKLPQ 291
QY 280 RMVRYLMSREITLGRATKQNDQIDVDLSLEGPAWKISKQGVITKKNNG 328
D 292 QVFOENIVKQDVQQLYRTGLPQ-SYVNAFAGDATKLDL---IYELKENG 336

RESULT 6
Tpr homolog - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C/Accession: T13829
R/Zimovska, G.; Aris, J.P.; Paddy, M.R.
J. Cell Sci. 110, 927-944, 1997
A/Title: A Drosophila Tpr protein homolog is localized both in the extrachromosomal char
A/Reference number: 217786; MUID:97296455; PMID:9152019
A/Accession: T13829
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-2346 <ZIM>
C/Cross-references: EMBL:U91980; NID:91923273; PID:91923274; PIDN:AA047506.1
C/Genetics:
A/Cross-references: FlyBase:FBgn0013756
A/Map position: 2R

Query Match 5.8%; Score 116; DB 2; Length 2346;
Best Local Similarity 23.3%; Pred. No. 13;
Matches 86; Conservative 55; Mismatches 126; Indels 102; Gaps 21;

QY 45 KRVKSKQPLQVTKDGRMKRANLILLINAVLQTNLITSVHLGVKFSCTFLREVGGRWY 104
D 340 KKIQMEQELASANDL-LKQARESNLSAICQLAPSAV-----ASLIRSDSLTEL 391
QY 105 ALLVDPVSKLACQAMQQLHPEAIAIOSKALPSKAEBO--LISKVGSQPLTFQDLL 162
D 392 YSMYAKSSEEL--EMRNCETQOL-KLQKSLIAISBAPFLBKQNSPYQMKETNSL 447
QY 163 LHRHDAFY-----LARTAKAL-----QAHMQLMQLKQYLLLEDQ----- 195
D 448 LREHDELLQNKLCLERELERALSTLNHQNENKQLQHTDLSRQVCMILDELINCRAGV 507
QY 196 ---TVQP---LPK-----GDQVLFNSDAEDLID-DSKLKMDDEVLEHLMVADR----- 239
D 508 KHVRIQTPRQQLPTSSLSIDNLVTFSSIEELVDRNTYLLNMSRELT--ELLEASERKNDK 565
QY 240 -----QKEIRIQLEQELHKQVAVLVSITGSSPDPNQTAVLRGRWY--LMSREIT 292
D 566 MLTBQSKHIRKLDK--RFALEDDLITQ-----KNTVTTLKSKCRKYLKYPAAQK 616
QY 293 LGRATKD-----NQIDVDLSLEGPA--WKISR-----QGVIK---LKN 326
D 617 LGQNTVDDDDSNLEPNDSLDTS-EQPAANFEKSRKLEKRVQLRQLEGVEVKYASLKE 675

```

```

QY 327 NGDFPIANE 335
D 676 NYDYTSEK 684

RESULT 7
Kinesin-related protein KRP85 - sea urchin (Strongylocentrotus purpuratus)
S38982
N/Alternate names: Kinesin-2 chain A; KRP (85/95) 85K chain
C/Species: Strongylocentrotus purpuratus (purple urchin)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C/Accession: S38982; S72551
R/Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.
Nature 366, 268-270, 1993
A/Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.
A/Reference number: S38982; MUID:94050179; PMID:8232586
A/Accession: S72551
A/Molecule type: mRNA
A/Residues: 1-699 <COL1>
A/Cross-references: EMBL:L16993; NID:9295245; PIDN:AAA16098.1; PID:9295246
A/Accession: S72551
A/Molecule type: protein
A/Residues: 2-5, 'X', '7-11; 59-64; 125-132; 222-226, 'X', 228-230 <COL2>
C/Complex: heterotrimer of a 115K chain and two kinesin-related chains of 95K (PIR:S5869
C/Superfamily: kinesin-related protein KIF3; kinesin motor domain homology
C/Keywords: Arp; heterotrimer; microtubule binding; nucleotide binding; P-loop
F;97-104/Region: nucleotide-binding motif A (P-loop)
F;103/Binding site: ATP (Lys) #status predicted

Query Match 5.8%; Score 115; DB 1; Length 699;
Best Local Similarity 20.6%; Pred. No. 2.8;
Matches 82; Conservative 58; Mismatches 138; Indels 120; Gaps 16;

QY 6 GVEPGRCGSGSPSSSE-----KKVSKAPSPVPPSPAPAGLTGRVYKSKQPLQV 57
D 377 GLDDDESGSESGDEAGGVKKRKGKPKRKL--SPETMAAQKIDBEKALEK 434
QY 58 KDLGRMKRANLILLINAVLQTNLITSVHLGVKFSCTFLREVGGRWYALLVDPVSKLAC 117
D 435 KD-----MVEDRNTVA-----RELQRR-----ESELHK--A 459
QY 118 QAMROLHPEAIAIOSKA-----LFSKAEQQLSKVGSQPLTFQDLLHRHPDAFY 171
D 460 QDDQKILNEKLAIOKLIYGVGVLLAKSBO-----EQLLIQ----- 497
QY 172 LARTAKALQAMQQLKQYLLLEDQTVQPLPKGQDVLFNSDAEDLID-DSKLKMDREVL 229
D 498 -----SALEMKERMAKQSMKRMKEERQERMDIEKYSLSQDEAHGKTKKLKQVWMLM 552
QY 230 EHELMVADRROKRE-----IRQLEQELHKQVAVLVSITGSSPDP-----NQT 274
D 553 QAKSEVADMOAHOHEEMALLEENVAELRELRSLMLIDISFIPQEFQIMIDQYVMMNDI 612
QY 275 AVLGRWYRYLMSREITLGRATKQNOI---DVDS-----LEGPAWKISRQGVITL 324
D 613 GEMQLKCAVYIGNNRKQTPVADKDKSLAYGEADLSNVFLTYNLBGGKMKTPSQG----- 668
QY 325 KNGDPFIANEGRPYIDGRPVVLGSKRWRLSNNSVVE 362
D 669 -----KSGRPXTSSGRP-KTGGKKQASMASSID 695

RESULT 8
A96494
protein F7P22.16 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: A96494
R/Theologas, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000

```

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Author: Salzer, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
ArTitle: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: A96494  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-804 <STO>  
A:Cross-references: GB:AE005173; NID:g6691192; PIDN:AAR24530.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F7F22.16  
A:Map position: 1

Query Match 5.6%; Score 111; DB 2; Length 804;  
Best Local Similarity 23.0%; Pred. No. 6.4; Indels 60; Gaps 11;  
Matches 67; Conservative 41; Mismatches 123;

Oy 97 REVERNYAL---LYDPVISKLACQAMROLH-PEAIIAIQSALFSKAEOLLSKVGS TS 152  
Db 47 RMATERREILDEVEDEMEPQTRATKLHKPDLLPTEBYRLFLNLFCSTRYCGSTS 106  
Oy 153 QPTLETFODLHHHPDAFYLRATKALOHWOLMKOYYLLEDQTVLPKGDVOLFNSDA 212  
Db 107 LAQLGLEDEVQH-----LYQSCHDLTLMAY-----PYVAVEBETLOPLSTGLVKYLXQGMT 156  
Oy 213 EDLIIDSKLMDRMREVLEHELMVADR-----QKEIREOLEGELHKWQ 255  
Db 157 SDELDCGEGLFGRFSVGHERLSTIKLEGLFGPSGTSGSKPKYEERBLKD L-----WT 210  
Oy 256 VLVDISIRMGSSPPDNQTLAVLRGRMRYLYMR-----SREITLGRTAKNDIQDVDSL 308  
Db 211 TTGGSVPLANSRSKSNC-----IRSPVIRYFQRSVANVLXSRETT-GTYVNSDMEMTAMAL 265  
Oy 309 EGPAWKISRKGQVIKLNNNGDFFIANEGRRPIYIDGRPVLCSK-WRLSN 358  
Db 266 -----KGTLRQTNGMSLQGKNNTPLSILLHLHCYGKSWAVSN 306

RESULT 9  
protein ZK742.1 [imported] - Caenorhabditis elegans  
G89113  
C.Species: Caenorhabditis elegans  
C.Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: G89113  
R.anonymus, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:8851916  
A>Note: see website genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Accession: G89113  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1140 <STO>  
A:Cross-references: GB:chr\_V; PIDN:AAB04981.1; PID:g1465818; GSPDB:GN00023; CESP:ZK742.1.L.1  
A:Map position: 5  
C:Superfamily: Arabidopsis thaliana exportin 1

Query Match 5.3%; Score 106.5; DB 2; Length 1140;  
Best Local Similarity 21.4%; Pred. No. 22;  
Matches 79; Conservative 61; Mismatches 155; Indels 75; Gaps 18;

Oy 50 SKQPLQVT-KDIGR---WKPANDLLIN--AVIQTNDLT-SVHLGVKFSGRFTLREVQR 102  
Db 217 SNSNQATLKTIQLPFTIPIGYVFETITELLSENFLSLVYRVVALQCLTEISQIVE 276  
Oy 103 WYALLDYPVSIKLACQAMROLHPEAIIAIQSALFSKAEOLLSKVGS TSQPTLETFODL 162

```

Db 277 TNDPSDEKLVKFCSGMRHISQVLSDLDLAAYYKASDDQKISSIAGQVLAVFIKSH 336
QY 163 LHRHPDAFYLTAKAL--QAH---WQIMKQYLLDQTVQPLPKGDQVLFNSDAED--- 214
Db 337 VHILVETDEPLTEAKILMRSHDYAIOILLKITLEEMEV-----FKVCLDWC 385
QY 215 -----LDDSKLKMRRDEVLEHMLVADRQ--KREIROLELHKQYLVDSI 261
Db 386 WLTAELVRLCPFIQPSFTLYGMSQVREH---PRRQLYREYLSQLRSTM-----I 431
QY 262 TGMSSPDPDNOTLAVL--RGRMYVLYLMSRR--ITLGRATKQNOIDVDLSLEGPAWKISRK 318
Db 432 SRRAKPE---EVLLVENDQGEVAVREWKDTSIALYRMRTETLYVL--THLDKDKTEVKKMT 487
QY 319 QGVYIKLKNNGDPFIANEGRRPYIDGRPVLCGSKRRS--NNSVVEIASLRFVFLINQDL 376
Db 488 EKLSQVNGGEFSWKINLR-----LC---WAVGSIQGTWEDEKRFVLVIVIRDL 534
QY 377 IALIRAEAAK 386
Db 535 LGLCEQKRGK 544

RESULT 10
G02520
C:Species: Homo sapiens (man)
C:Accession: G02520 #sequence_revision 06-Jun-1997 #text_change 20-Aug-1999
C:Accession: G02520
R:McLean, W.H.I., Smith, F.J.D.
submitted to the EMBL Data Library, March 1996
A:Reference number: H01385
A:Accession: G02520
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4574 <MCU>
A:Cross-references: EMBL:U53204; NID:g1477645; PIDN:AA05427.1; PID:g1477646
C:Gene: PLECL
C:Genetics:
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S
F:68-283/Domain: alpha-actinin actin-binding domain homology <ACT>

Query Match 5.3%; Score 106.5; DB 2; Length 4574;
Best Local Similarity 23.5%; Pred. No. 1.5e+02;
Matches 70; Conservative 52; Mismatches 119; Indels 57; Gaps 13;

QY 122 QLMPEA-IAIISKALPSKAEOLSKVSGTSOPTLETFFODLHRHPDAFYLTAKALQ 180
Db 2488 KIQQEAQLQLKSEBQTVQOEQLQETQALQOSTLSKDSILQK--ERFIQERAKLEQ 2545
QY 181 AHW-QLMKQYLLDQTVQPLPKGDQVLFNSDAEDLIDSKLKMDRDEVLEHMLVADR 239
Db 2546 LFDQVAKAQAQLNEEQQRQ-----QQQMEGERQLV--ASWEARRR--QHEAEGVRR 2595
QY 240 QKREIROLELHKQYLVDSITGMSPPDPNQ---TLAVLRGMVRYLMSRRETTLGR 295
Db 2596 KQRELQOLEQRRQOEELAE-----ENQRLREQLOLEEQRALAHSEVETASQ 2646
QY 296 --ATKQNOIVDSLSEPAWK-----ISRKGQVILKNNGDFTANEGRRPI---- 340
Db 2647 VAAATKTLPNGRD-ALDSDPAABEPHSEFDGIRRKYSQAQRLQAG--ILSHBELQRLAQGHT 2704
QY 341 -----YIDRPPVLCGSKRWLSNNSVVEIASLRFVFLINQDLIALIRAEAA 385
Db 2705 TVDELARREDVRYLQGRSSISAGILLKATNEKLSVYALALQQLSPGVALILLDAQAA 2762

RESULT 11
A59404
Plectin [imported] - human
C:Species: Homo sapiens (man)
C:Accession: A59404 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
C:Accession: C59404

```



R./Iiu, C.G.; Maercker, C.; Castanon, M.J.; Hauptmann, R.; Wiche, G.  
 Proc. Natl. Acad. Sci. U.S.A. 93, 4278-4283, 1996  
 A/Title: Human plectin: organization of the gene, sequence analysis, and chromosome loca  
 A/Reference number: C59404; MUID:96210632; PMID:8633055  
 A/Accession: C59404  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-4684 <STO>  
 A/Cross-references: GB:CAA91196; NID:g1296662; PIDN:CAA91196.1  
 C/Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S

Query Match 5.3%; Score 106.5; DB 2; Length 4684;  
 Best Local Similarity 23.5%; Pred. No. 1.66+02;  
 Matches 70; Conservative 52; Mismatches 119; Indels 57; Gaps 13;

Qy 122 QLHPEA-IALAQSRLFKAEHQLLSKYGSTQPTLFETFDLHHPDAFYIARAKLQ 180  
 Db 2598 KLQOEAKLQKSEMQVVOEQLLQETQALQOSFLSEKDSLQF--BRFIEQEKAKLEQ 2655  
 Qy 181 AHM-QLMQVYLLEPQVPLPKQDVLFNSDAEDLIDDSKDKMRDEVLHEHLMVADR 239  
 Db 2656 LFODEVAKAQQUREQQRQ-----QQQMEQERQRLV--ASNEEARRR--QHAESEGVRR 2705  
 Qy 240 QKREIRQLEQELHKQVVDSTGMSPPDFNQ---TLAVLRGMVRYLMSREITLGR 295  
 Db 2706 KQELQLEQQRQOEELAE-----ENQRLREQQLLEQHRALAHSEEVTAQ 2756  
 Qy 296 --ATKQNDIVDLSEGRAMK-----ISRKQGVIKLKNQPFIFIANEGRRI---- 340  
 Db 2757 VAATKTLPLNGRD-ALDGPAAEPEPESFDGKRKRSQAORLEAG--ILSAEELQRIAQCHT 2814  
 Qy 341 -----YIDGPRVLGSGKMRISNNNSVVEIASRFVFLINODIALIRAEAA 385  
 Db 2815 TVDELAREDEVRAHYLGKSSISGLIKATNEKLSVYALQRLSPGTRALLILEKQAA 2872

RESULT 12  
 A27605  
 N/Alternate names: Duchenne muscular dystrophy protein  
 C/Species: Homo sapiens (man)  
 C/Date: 19-Nov-1998 #sequence revision 27-Jun-1994 #text change 16-Jun-2000  
 C/Accession: A27605; S07710; A27162; S05291; A40134; S06051; S10346; S02242; S02  
 R/Koenig, M.; Monaco, A.P.; Kunkel, L.M.  
 Cell 53, 219-228, 1988  
 A/Title: The complete sequence of Duchetrophin predicts a rod-shaped cytoskeletal protein.  
 A/Reference number: A27605; MUID:88194521; PMID:3282674  
 A/Accession: A27605  
 A/Molecule type: mRNA  
 A/Residues: 1-3685 <KOE>  
 A/Cross-references: GB:M18533; NID:g181856; PIDN:MAA53189.1; PID:g181857  
 R/Rosenthal, A.; Speer, A.; Billwiltz, H.; Cross, G.S.; Forrest, S.M.; Davies, K.E.  
 Nucleic Acids Res. 17, 5391, 1989  
 A/Title: Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus  
 A/Reference number: S07710; MUID:89345106; PMID:266885  
 A/Accession: S07710  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-132; 'P', 134-622, 'I', 624-783, 'G', 785-1196, 'F', 1198-1376, 'N', 1378-1468, 'Q', 1  
 A/Cross-references: EMBL:X14298; NID:g30845; PIDN:CAA3479.1; PID:g30846  
 A/Note: this sequence was submitted to the EMBL Data Library, February 1989  
 R/Koenig, M.; Hoffman, E.P.; Bertelson, C.J.; Monaco, A.P.; Feener, C.; Kunkel, L.M.  
 Cell 50, 509-517, 1987  
 A/Title: Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary  
 A/Reference number: A90897; MUID:87273512; PMID:3607877  
 A/Accession: A27162  
 A/Molecule type: mRNA  
 A/Residues: 1-497 <KO2>  
 A/Cross-references: GB:M18533  
 R/Cross, G.S.; Speer, A.; Rosenthal, A.; Forrest, S.M.; Smith, T.J.; Edwards, Y.; Flint,  
 EMBO J. 6, 3277-3283, 1987  
 A/Title: Deletions of fetal and adult muscle cDNA in Duchenne and Becker muscular dystro  
 A/Reference number: S01263; MUID:8811512; PMID:3428261

A/Accession: S05291  
 A/Molecule type: mRNA  
 A/Residues: 404-556; 'T', 558-610, 'K', 612-622, 'I', 624-664, 'M', 665-783, 'G', 785-1137, 'PN' <C  
 A/Cross-references: EMBL:X06178  
 A/Note: 475-Ile and 529-Glu were also found  
 R/Hoffman, E.P.; Monaco, A.P.; Feener, C.C.; Kunkel, L.M.  
 Science 238, 347-350, 1987  
 A/Title: Conservation of the Duchenne muscular dystrophy gene in mice and humans.  
 A/Reference number: A40134; MUID:88018015; PMID:3659917  
 A/Accession: A40134  
 A/Molecule type: mRNA  
 A/Residues: 491-1207 <HOF>  
 A/Cross-references: GB:M18533  
 R/Blonden, L.A.J.; den Dunnen, J.T.; van Paassen, H.M.B.; Wapenaar, M.C.; Grootscholten,  
 Nucleic Acids Res. 17, 5611-5621, 1989  
 A/Title: High resolution deletion breakpoint mapping in the DMD gene by whole cosmid hyb  
 A/Reference number: S06051; MUID:8934515; PMID:2569720  
 A/Accession: S06051  
 A/Status: translation not shown  
 A/Molecule type: DNA  
 A/Residues: 2147-2204 <ELO>  
 A/Cross-references: EMBL:X15495; NID:g30829; PIDN:CAA33518.1; PID:g1335049  
 R/Speer, A.; Billwiltz, H.; Hutch, A.; Couelle, C.; England, S.; Love, D.; Davies, K.E.  
 submitted to the EMBL Data Library, February 1990  
 A/Reference number: S10346  
 A/Accession: S10346  
 A/Molecule type: DNA  
 A/Residues: 2438-2480 <SPE>  
 A/Cross-references: EMBL:X51934  
 R/Chamberlain, J.S.; Gibbs, R.A.; Ranier, J.E.; Nguyen, P.N.; Caskey, C.T.  
 Nucleic Acids Res. 16, 11141-11156, 1988  
 A/Title: Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA a  
 A/Reference number: S02109; MUID:8908552; PMID:3205741  
 A/Accession: S02243  
 A/Status: preliminary; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 665-722 <CHA>  
 A/Cross-references: EMBL:X13045; NID:g30825; PIDN:CAA31451.1; PID:g1335048  
 A/Accession: S02242  
 A/Status: preliminary; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 2098-2146 <CH2>  
 A/Cross-references: EMBL:X13047; NID:g30827; PIDN:CAA31452.1; PID:g809549  
 A/Accession: S02244  
 A/Status: preliminary; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 2147-2204 <CH3>  
 A/Cross-references: EMBL:X13048; NID:g30833; PIDN:CAA31454.1; PID:g1335051  
 A/Accession: S02109  
 A/Status: preliminary; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 2305-2365, 'K' <CH4>  
 A/Cross-references: EMBL:X13047; NID:g30831; PIDN:CAA31453.1; PID:g1335050  
 R/Giljaar, I.H.B.; van Paassen, M.H.M.B.; den Dunnen, J.T.; van Ommen, G.G.J.B.  
 submitted to the EMBL Data Library, March 1992  
 A/Description: Sequence of Duchenne muscular dystrophy gene exon 60, located directly 5'  
 A/Reference number: S23736  
 A/Accession: S23736  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 2980-2995, 'K', 2997-3028 <GIN>  
 A/Cross-references: EMBL:Z11860  
 R/Hellig, R.; Lemaire, C.; Mandel, J.L.  
 Nucleic Acids Res. 15, 9129-9142, 1987  
 A/Title: A 230bp cosmid walk in the Duchenne muscular dystrophy gene: detection of a con  
 A/Reference number: S09071; MUID:88067745; PMID:2825128  
 A/Accession: S09071  
 A/Molecule type: DNA  
 A/Residues: SGGGSHWTHGSLIYRLPLTL', 218-277 <HET>  
 A/Cross-references: EMBL:X06293; EMBL:Y00494  
 A/Note: sequence N-terminal of residue 218 correspond to a putative exon  
 R/Roberts, R.G.; Coffey, A.J.; Bobrow, M.; Bentley, D.R.  
 Genomics 16, 536-538, 1993

A:Title: Exon structure of the human dystrophin gene.  
A:Reference number: I54186; MUID:93300536; PMID:831593  
A:Accession: I54186  
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL  
A:Molecule type: DNA  
A:Residues: 984-1411 <RE3>  
A:Cross-references: GB:I05642; NID:G181892; PIDN:AAA74506.1; PID:G950344  
A:Accession: I68509  
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL  
A:Molecule type: DNA  
A:Residues: 1776-1913 <RE2>  
A:Cross-references: GB:I05646; NID:G181896; PIDN:AAA74507.1; PID:G950345  
A:Accession: I68510  
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL  
A:Molecule type: DNA  
A:Residues: 2850-2979 <ROB>  
A:Cross-references: GB:I05649; NID:G181899; PIDN:AAA74508.1; PID:G950346  
A:Robert, R.G.; Coffey, A.U.; Bobrow, M.; Bentley, D.R.  
Genomic 13, 942-950, 1992  
A:Title: Determination of the exon structure of the distal portion of the dystrophin gene  
A:Reference number: I54175; MUID:92372062; PMID:1505985  
A:Accession: I54175  
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL  
A:Molecule type: DNA  
A:Residues: 2980-3685 <RES>  
A:Cross-references: GB:M86903; NID:G181881; PIDN:AAA35779.1; PID:G457519  
A:Benatre, J.; Hillier, M.; Junkes, B.; Piord, M.; Schwinger, E.; Voeberg, H.P.  
Genomic 10, 551-557, 1991  
A:Title: Analysis of a dystrophin gene deletion by amplification of mRNA isolated from D  
A:Reference number: I54166; MUID:91365360; PMID:1889805  
A:Accession: I54166  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 2250-2254 <RE4>  
A:Cross-references: GB:S54699; NID:G235303; PIDN:AA19754.1; PID:G235304  
A:Feener, C.A.; Koenig, M.; Kunkel, L.M.  
Nature 338, 509-511, 1989  
A:Title: Alternative splicing of human dystrophin mRNA generates isoforms at the carboxy  
A:Reference number: S03902; MUID:89181947; PMID:2648158  
A:Accession: S03902  
A:Molecule type: mRNA  
A:Residues: "MED", 12-32/3377-3408 <PEE>  
A:Cross-references: EMBL:X15148  
C:Comment: Dystrophin is proposed to play a role in anchoring the cytoskeleton to the pl  
C:Comment: Defects in dystrophin are responsible for the Duchenne/Becker muscular dystro  
C:Keywords: actin binding; alternative splicing; calmodulin binding; cytoskeleton; leuc  
elix  
F:114-233/Domain: alpha-actinin actin-binding domain homology <ACT>  
F:253-327/Region: hinge  
F:338-447/Domain: spectrin/dystrophin repeat homology <SP01>  
F:448-556/Domain: spectrin/dystrophin repeat homology <SP02>  
F:558-667/Domain: spectrin/dystrophin repeat homology <SP03>  
F:668-717/Region: hinge  
F:718-828/Domain: spectrin/dystrophin repeat homology <SP04>  
F:838-934/Domain: spectrin/dystrophin repeat homology <SP05>  
F:938-1045/Domain: spectrin/dystrophin repeat homology <SP06>  
F:1047-1154/Domain: spectrin/dystrophin repeat homology <SP07>  
F:1156-1263/Domain: spectrin/dystrophin repeat homology <SP08>  
F:1265-1367/Domain: spectrin/dystrophin repeat homology <SP09>  
F:1372-1477/Domain: spectrin/dystrophin repeat homology <SP10>  
F:1478-1568/Domain: spectrin/dystrophin repeat homology #atypical <SP11>  
F:1570-1676/Domain: spectrin/dystrophin repeat homology <SP12>  
F:1678-1782/Domain: spectrin/dystrophin repeat homology <SP13>  
F:1784-1875/Domain: spectrin/dystrophin repeat homology #status atypical <SP14>  
F:1876-1982/Domain: spectrin/dystrophin repeat homology <SP15>

F:1984-2101/Domain: spectrin/dystrophin repeat homology <SP16>  
F:2103-2208/Domain: spectrin/dystrophin repeat homology <SP17>  
F:2210-2316/Domain: spectrin/dystrophin repeat homology <SP18>  
F:2317-2423/Domain: spectrin/dystrophin repeat homology <SP19>  
F:2424-2470/Region: hinge  
F:2471-2577/Domain: spectrin/dystrophin repeat homology <SP20>  
F:2579-2686/Domain: spectrin/dystrophin repeat homology <SP21>  
F:2688-2802/Domain: spectrin/dystrophin repeat homology <SP22>  
F:2804-2931/Domain: spectrin/dystrophin repeat homology <SP23>  
F:2933-3040/Domain: spectrin/dystrophin repeat homology <SP24>  
F:3041-3112/Region: hinge  
F:3055-3092/Domain: WW repeat homology <WW1>  
F:3080-3360/Region: cysteine-rich  
F:3506-3527/Region: leucine zipper motif  
F:3572-3593/Region: leucine zipper motif  
Query Match 5.3%; Score 106; DB 1; Length 3685;  
Best Local Similarity 21.2%; Pred. No. 1.2e+02;  
Matches 70; Conservative 59; Mismatches 121; Indels 80; Gaps 14;  
QY 32 PVPSPAPAPGCTKTVKSKOPLOVTKGRKRPANDLLITNAVLOTNDLVHIGVKFS 91  
DB 3417 PVDASAPASSPOLSHDTHSRLEHYASR-LAEMENSGSYLNDISPNSIDDEHLLIHY 3475  
QY 92 CEFTRREVQERVYALLYPVISKLACQMRQLHPAIAIQSKALFKAEOILSKVST 151  
DB 3476 CQ-----SLNQSPLSQ-----PSPAQILISLSE 3501  
QY 152 SQPTLE-TFQDLLHRHPAFYLARTAKALQAHWQMLKQVYLLEDVTPPLPKGDVLNFS 210  
DB 3502 ERGELERLIADLEEN-----RNLQARYDRLKQOH--EHKGLSPLPSPPMMPFS 3549  
QY 211 -----DAEDLIDSKLKMREDVLEHELMVADNRQKREIRQLQELHKQVYVD----- 259  
DB 3550 POSPPDAR-LIAEALRLROHKGRLEARQQLIEDHNK---QLSEQLHRRLQLEPPQAEA 3604  
QY 260 --SITGSSPPDQNTLAVLRGMVRYLMRSREITLGRATKQNDVDSLSEGPAMKISR 317  
DB 3605 KNGTIVSSPFSLDRSSSQPMILR-----VGSQTSBSMGEDFL-LSPPDITSG 3655  
QY 318 KQGVTKNNNGDFIANEGRRPIYIDGRPV 347  
DB 3656 LEEVMEQLNNS--FPSSGRN---TPGKPM 3680  
RESULT 13  
T28667  
dynein beta heavy chain - Paramesitum tetraurelia  
C:Species: Paramesitum tetraurelia  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Mar-2001  
C:Accession: T28667; PC4340  
R:Kandl, K.A.; Forney, J.D.; Asai, D.J.  
submitted to the EMBL Data Library, January 1995  
A:Description: The dynein genes of Paramesitum: the differential expression of axonemal a  
A:Reference number: Z20502  
A:Accession: T28667  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4588 <KAN>  
A:Cross-references: EMBL:U19464; NID:G625089; PID:G625090; PIDN:AAA61680.1  
R:Asai, D.J.; Beckwith, S.M.; Kandl, K.A.; Keating, H.H.; Tjandra, H.; Forney, J.D.  
J. Cell Sci. 107, 839-847, 1994  
A:Title: The dynein genes of Paramesitum tetraurelia: Sequences adjacent to the catalytic  
A:Reference number: PC4340; MUID:94334383; PMID:8036840  
A:Accession: PC4340  
A:Molecule type: mRNA  
A:Residues: 1886-2085 <ASA>  
C:Keywords: dynein; dynein heavy chain, ciliary  
C:Keywords: nucleotide binding; P-loop  
F:1886-1893/Region: nucleotide-binding motif A (P-loop)

Query Match 5.2%; Score 104.5; DB 2; Length 4588;  
Best Local Similarity 21.5%; Pred. No. 2.1e+02;  
Matches 59; Conservative 58; Mismatches 72; Indels 91; Gaps 17;

Qy 48 KKSQPLQVTKDGLGRKRPANDLLINAVLQTNDLTSVHLGVKFSGRFTLREVERVALL 107  
Db 650 EKQKQPL-LSKD-----ENGLLRVNF-----DPALVRL-----LREV--KYFTLL 686  
Qy 108 YDPVTSKLACQAMQMLHEALAIQSKALFSAKEBQLSKVSGTSQPLTFEFDLHHP 167  
Db 687 EQPV-----PESASELYSKN-----DTFRE----- 706

Qy 168 DAFYLARTAKALQAHQMLQKQVYLLLEDQTVQ-PLPKGPQVL-----NPSDAEDLID 217  
Db 707 ---YIVQEMIVENNFIYVTLQHPHEPLIKRIEKMDEVLKPGIEHKYKMSNDINKETIE 763

Qy 218 DSKLDMDEVLEHMLVADRROKREIRQLDEQELHKW--QVLVDSITGMSPPDPDQTLA 275  
Db 764 TAKA--TVDEL--HQIV---QMKETLKKIEQALEKFTMTKIERNKXKMPSPDYQFLKA 816

Qy 276 VLGRMVRVYLMRSREITIGRATKD--NOIDVDLSLEGPAW 313  
Db 817 VVQNKV--SIVKNGTSTINKLVKEVLQVKKVKKQO--AW 852

RESULT 14  
G96522  
F11A17.16 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: G96522  
R/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
R/Author: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marzali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: G96522  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-558 <STO>  
C/Genetics:  
A/Cross-references: GB:A805173; NID:g8778962; PIDN:ADD49768.2; GSPDB:GN00141  
A/Map position: 1

Query Match 5.2%; Score 103.5; DB 2; Length 558;  
Best Local Similarity 21.5%; Pred. No. 13;  
Matches 88; Conservative 51; Mismatches 129; Indels 141; Gaps 18;

Qy 7 VEPGRSGSEPSSS-----EKKVSKASTVPPSPAPAP-GL 43  
Db 214 VESSRLSPSPSPSRPLPTPLPKFLVPASSLGRDENSPPFPPTPPPPPPPPPL 273

Qy 44 TKRVKSKQPLQVTKDGLGRKRPANDL-LINAVLQTNDLTSVHLGVKFSGRFTLREVER 102  
Db 274 AKAAQKQSP-----FVSQFLQNLKQDNRSLISQSVNKKQVNSA----- 315

Qy 103 WYALIDPVISKLACQAMQMLHEALAIQSKALFSAKEBQLSKVSGTSQPLTFEFD 161  
Db 316 -----HNSIVGEI---QRSALILAIKADIERKEFI---NDLIQKVLTTC----- 355

Qy 162 LILHHPDAFYLTARTAKALQAHQMLKQYVLLLEDQTVQPLPKGDQVLNPSDAEDLID 217  
Db 356 -----FSDMEDVKKFVDW 368

Qy 218 -DSKLQKMRDE--VLEH---ELMVADRROK-----REIRQLDEQELHKWQVYLVDSITGMS 266

Db 369 LDKELATLADRAVLKPIFKWPEKKADTLQEAAYEYRELKLEKELSSYSDDPNIHYVAL 428  
Qy 267 PDPNQTLAVLRGMVRYLMSREITIGRATKQNOIDVDLSL-GRPAKISRKGIVTLK 325  
Db 429 KQMAN--LDDKSEQRIRRLVLRGSSM--RSYQDFKIPYEMWLDGMI CKI--KASITLA 483

Qy 326 NNGDFIANEGRPDIYDGRVLCGSKWRLSNNGSVETIASLRPFVLNQ 374  
Db 484 KTYMNRVANELQSANLD-----RESTKALLQGVRFAYRTHQ 522

RESULT 15  
T42725  
actin binding protein ACF7, neural isoform 1 - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 03-Nov-2000  
C/Accession: T42725  
R/Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.  
Genomic 38, 19-29, 1996  
A/Title: Cloning and characterization of mouse ACF7, a novel member of the dyxectonin subf.  
A/Reference number: Z20900; MUID:97124842; PMID:8954775  
A/Accession: T42725  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1825 <BER>  
A/Cross-references: EMBL:U67203; NID:g1675221; PID:g1675222; PIDN:ACC52988.1  
C/Genetics:  
A/Gene: ACF7  
A/Map position: 4  
C/Superfamily: plectrin; alpha-actinin actin-binding domain homology; ribosomal protein S  
C/Keywords: actin binding

Query Match 5.2%; Score 103.5; DB 2; Length 1825;  
Best Local Similarity 21.6%; Pred. No. 69;  
Matches 71; Conservative 59; Mismatches 124; Indels 75; Gaps 15;

Qy 52 QPLVTKDGLGRKRPANDLLINAVLQTNDLTSVHLGVKFSGRFTLREVO-----ERMYA 105  
Db 540 EEMQKLERAEW--GNDLPSVELQLETOQ--HHTSVF-ELGSSVKEARLYEGKMSQNFH 594

Qy 106 LLYDPVTSKLACQ-----AMRQLHPEALAIQSKA-----LFSKAEQQLSKVSGT 151  
Db 595 TSYVETLQKLETOYCKLKETSFFMRHQLSLKRVSRATLILNNGKEEBELACD--WSD 653

Qy 152 SQPT-----LETFQDL-----LHRHPDAFYLTARTAKALQAHQML 185  
Db 654 SNPNISAKKTYFSELTMELBGQDVFRSLQDTAEVLSLENHPAKQTVBAYSAVQSQLOW 713

Qy 186 MKQYVLLLEDQTVQPLPKGDQVLNPSDAEDLIDSKLDMDEVLEHMLVADRRO--KRE 243  
Db 714 MKQCLCVQGHKENVAYFQF--FSDARDL--ESFLRLQDSI--KKRYTDRSTLSRL 767

Qy 244 IROLEQLHKQVYLVDSITGMSPPDPDQTLAVLRGMVRYLMSR-----EI 291  
Db 768 EDLIQDSMDEKEQLIQSSVASLVSRSKTIYQLKPRNPDLVASTLSVKALCOYRQLEI 827

Qy 292 TLGRATKQNOIDVDLSLEGPAKISRKQG 320  
Db 828 TI---CKNDECVLEBDSQRTKMKVISPFG 853

Search completed: March 27, 2003, 05:22:44  
Job time : 57 secs

**This Page Blank (uspto)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 27, 2003, 05:06:28 ; Search time 23 Seconds  
(without alignments)

703.294 Million cell updates/sec

Title: US-09-744-125A-2

Perfect score: 1996  
Sequence: 1 NSARGVPEGRCSGSEPPSS.....LINODLIALIRAEAKITPQ 390

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	5.8	699	1 K122_STRPU	P6872 Strongyloce
2	106.5	5.3	4684	1 PLEI_HUMAN	Q15149 homo sapien
3	106	5.3	3685	1 DMD_HUMAN	P15132 homo sapien
4	103.5	5.2	5327	1 ACF7_MOUSE	O9qxx0 mus musculu
5	103	5.2	810	1 GYRA_BORBU	O51366 borrelia bu
6	102	5.1	3678	1 DMD_MOUSE	P15531 mus musculu
7	101.5	5.1	1109	1 POL_CAESC	P33459 caprine art
8	101	5.1	3680	1 DMD_CANFA	O97592 canis fam11
9	99.5	5.0	417	1 SYS_MYCSE	P47251 mycoplasma
10	99.5	5.0	496	1 H1A2_ARATH	O06402 arabidopsis
11	99.5	5.0	911	1 T2D1_DROME	P41123 drosophila
12	99.5	5.0	2068	1 P7075_AQUPE	P70755 aquilex pyr
13	99	5.0	2244	1 P7075_AQUPE	P70755 aquilex pyr
14	98.5	4.9	855	1 DYNH_HUMAN	O14203 homo sapien
15	98.5	4.9	4486	1 DYNH_HUMAN	O14203 homo sapien
16	98	4.9	1278	1 DYNH_HUMAN	O14203 homo sapien
17	98	4.9	1453	1 V373_BOVIN	O82u22 bos taurus
18	97.5	4.9	598	1 LEPA_MYCPN	P75498 mycoplasma
19	97.5	4.9	1048	1 SBCC_ECOLI	P33458 escherichia
20	97.5	4.9	4644	1 DYHC_MOUSE	O9jhu4 mus musculu
21	97.5	4.9	4644	1 DYHC_MOUSE	O9jhu4 mus musculu
22	97	4.9	1225	1 KFA4_CHICK	O90640 gallus gall
23	96	4.8	4442	1 ODO2_RAT	Q01205 rattus norv
24	96	4.8	3660	1 DMD_CHICK	P11533 gallus gall
25	95.5	4.8	503	1 AGP4_HUMAN	O9y264 homo sapien
26	95.5	4.8	935	1 KINH_SYNRA	O43093 syncephalas
27	95.5	4.8	5430	1 ACF7_HUMAN	O9upn3 homo sapien
28	95	4.8	870	1 CSX2_SCHPO	O9ue23 echinosacch
29	95	4.8	2842	1 APC_EAT	P70478 rattus norv
30	94.5	4.7	458	1 RAD4_HABIN	P45266 haemophilus
31	94.5	4.7	754	1 KATC_ARATH	P46875 arabidopsis
32	94.5	4.7	883	1 CLPB_SYNP7	P53533 synecococc
33	94.5	4.7	1033	1 Y328_MYCPN	P75310 mycoplasma

34	94.5	4.7	2469	1 TEGU_HSVSA	O01056 herpesvirus
35	94	4.7	986	1 GM13_RAT	O62839 rattus norv
36	94	4.7	1539	1 Y373_HUMAN	O15078 homo sapien
37	93.5	4.7	409	1 ODO2_FUGRU	O90512 fugu rubrip
38	93	4.7	339	1 RLAD_ARCFU	O28781 archaeglob
39	93	4.7	402	1 TAU_CAPHI	O02828 capra hircu
40	93	4.7	404	1 Y349_MYCSE	P47531 mycoplasma
41	93	4.7	447	1 TAU_BOVIN	O92842 bos taurus
42	93	4.7	452	1 Y190_CHLPN	O92842 chlamydia p
43	93	4.7	695	1 APP2_MOUSE	O06335 mus musculu
44	93	4.7	852	1 GCP2_DROME	O9xyp7 drosophila
45	92.5	4.6	578	1 LIPA_MYCPN	O50274 mycoplasma

## ALIGNMENTS

RESULT 1  
K122\_STRPU STANDARD; PRT; 699 AA.  
ID K122\_STRPU  
AC P6872;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Kinesin-II 85 kDa subunit (KRP-85/95 85 kDa subunit).  
GN KRP85.  
OS Strongylocentrotus purpuratus (Purple sea urchin).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
OC Echinoidae; Euechinoidae; Echinacea; Echinoidae; Strongylocentrotidae;  
OC Strongylocentrotus.  
OX NCBI\_TaxID=7668;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Bg;  
RX MEDLINE=94050179; PubMed=8232586;  
RA Cole D.G., Chinn S.W., Wedaman K.P., Hall K., Vuong T.,  
RA Scholey J.M.;  
RT "Novel heterotrimeric kinesin-related protein purified from sea  
RT urchin eggs";  
RL Nature 366:268-270(1993).  
CC -1- SUBUNIT: Heterotrimer of a 115 kDa subunit (KAP115) and two  
CC -1- kinesin-like subunits of 95 kDa (KRP95) and 85 kDa (KRP85).  
CC -1- PTM: THE N-TERMINUS IS BLOCKED.  
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN  
CC II SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: L16993; AAA16098.1; -  
CC HSSP: P17119; 3KAR  
DR InterPro: IPR001752; kinesin\_motor.  
DR Pfam: PF00225; kinesin\_1.  
DR PRINTS: PR00380; KINESINHEAVY.  
DR SMART: SM00129; KISC; 1.  
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; 1.  
DR PROSITE: PS50067; KINESIN MOTOR DOMAIN2; 1.  
KW Motor protein; Microtubules; ATP-binding; Coiled coil.  
FT DOMAIN 1 340 KINESIN-MOTOR (BY SIMILARITY).  
FT DOMAIN 341 619 COILED COIL (BY SIMILARITY).  
FT DOMAIN 620 699 GLOBULAR (BY SIMILARITY).  
FT NP BIND 97 104 ATP (POTENTIAL).  
SQ SEQUENCE 699 AA; 78697 MW; 7B386611C808190 CRC64;  
Query Match 5.8%; Score 115; DB 1; Length 699;  
Best Local Similarity 20.6%; Pred. No. 0.89;  
Matches 82; Conservative 58; Mismatches 138; Indels 120; Gaps 16;

```

QY 6 GVEPGRSGSPSSSE-----KKVSKAPSTPPSPAPAPGTLTKVKKSKOPLQVT 57
D 377 GIDDEESGSESSGDEEAGEGCVKKKKKPKRKU--SPETMAAQKKIDEKALAEK 434
QY 58 KDLGRWKRPANDLLINAVLOTNDLTSHLVGVKPSCFRLREVOERWYALLYPVISKLAC 117
D 435 KD-----MVEDRNTVH-----RELQRR-----ESELHK--A 459
QY 118 QAMROLHPBALAIAISKA-----LFSKAEBOILSKVSTOPTLETFODLLHRHPDAFY 171
D 460 QDDQKILNEKNAIOKKLIVGVGDVLAKSEEO-----EOLLEQ----- 497
QY 172 LARTKALQAHQMLKQOYVYLEDQYVQPLPKGDQVYLFPSDADLI--DDSKLKDREVL 229
D 498 -----SALEMERRKAKQESMRKMRMEBOERMDIEBKYSLODEAHGKTKKLKVWYMLM 552
QY 230 EHELVADRQRQRE-----IRQLEQLHKQOVLVDSTGMSPPDF-----NQT 274
D 553 QAKSEVADMQAHQREMEALLENVRELRLSLMLIDSPFOERQEMI EQYVHNMEDI 612
QY 275 AVLGRMRYLYMRSEBITLGATKCNQI--DVDS-----LGGPAWKIRKGVYKL 324
D 613 GEMQKCVAYTCNNMRKQTPVADKDKSLAYGADLSNVFLTYNLBGGKVKYPSQG---- 668
QY 325 KNGDFFLANEGRPDIYIDGRPVLCGSKWRLSNNSGYE 362
D 669 -----KSGRPKTSGRP-KTGKKKQASMASSID 695

RESULT 2
ID_PLE1_HUMAN STANDARD; PRT: 4684 AA.
AC Q15149; Q16640; Q15148;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1).
GN PLECL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=96210632; PubMed=8633055;
RA Liu C.-G., Maercker C., Castanon M.J., Hauptmann R., Wiche G.;
RT "Human plectin: organization of the gene, sequence analysis, and
RT chromosome localization (8q24).";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283(1996).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND DISEASE.
RX MEDLINE=96312447; PubMed=8698233;
RA McLean W.H.I., Pulkkinen L., Smith F.J.D., Rugg E.L., Lane E.B.,
RA Bullrich F., Burgess R.E., Amano S., Hudson D.L., Owaribe K.,
RA McGrath J.A., McMillan J.R., Eady R.A.J., Leigh I.M., Cristiano A.M.,
RA Uitto J.;
RT "Loss of plectin causes epidermolysis bullosa with muscular dystrophy:
RT cDNA cloning and genomic organization.";
RL Genes Dev. 10:1724-1735(1996).
RN [3]
RP VARIANT MD-EBS 1003-GLN--ALA-1005 DEL.
RX MEDLINE=97049959; PubMed=8894687;
RA Pulkkinen L., Smith F.J.D., Shmizu H., Murata S., Yacita H.,
RA Hachisaka H., Nishikawa T., McLean W.H.I., Uitto J.;
RT "Homozygous deletion mutations in the plectin gene (PLECL1) in patients
RT with epidermolysis bullosa simplex associated with late-onset
RT muscular dystrophy.";
RL Hum. Mol. Genet. 5:1539-1546(1996).
RN [4]
RP VARIANT MD-EBS LEU-429 INS.
RX MEDLINE=21090821; PubMed=1159198;
RA Bauer J.W., Rouan F., Kofler B., Reniczek G.A., Kornacker I.,

```

```

RA Mus W., Hametner R., Klausegger A., Huber A., Pohl-Gubo G.,
RA Wiche G., Uitto J., Hantner H.;
RT "A compound heterozygous one amino-acid insertion/nonsense mutation in
RT the plectin gene causes epidermolysis bullosa simplex with plectin
RT deficiency.";
RL Am. J. Pathol. 158:617-625(2001).
CC -1- FUNCTION: INTERLINGS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND
CC MICROFILAMENTS AND ANCHORS INTERMEDIATE FILAMENTS TO DESMOSES OR
CC HEMIDESMOSES. COULD ALSO BIND MUSCLE PROTEINS SUCH AS ACTIN TO
CC MEMBRANE COMPLEXES IN MUSCLE. MAY BE INVOLVED NOT ONLY IN THE
CC CROSSLINKING AND STABILIZATION OF CYTOSKELETAL INTERMEDIATE
CC FILAMENTS NETWORK, BUT ALSO IN THE REGULATION OF THEIR DYNAMICS.
CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST LEVELS IN
CC MUSCLE, HEART, PLACENTA AND SPINAL CORD.
CC -1- DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN, THE C-TERMINUS WITH
CC VIMENTIN, DESMIN, GAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N-
CC AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.
CC -1- PTM: PHOSPHORYLATED BY CDC2, REGULATES DISSOCIATION FROM
CC INTERMEDIATE FILAMENTS DURING MITOSIS (BY SIMILARITY).
CC -1- DISEASE: DEFECTS IN PLECL1 ARE THE CAUSE OF EPIDERMOLYSIS BULLOSA
CC WITH MUSCULAR DYSTROPHY (MD-EBS OR EB-MD), AN AUTOSOMAL RECESSIVE
CC DISORDER CHARACTERIZED BY EPIDERMAL BLISTER FORMATION AT THE LEVEL
CC OF THE HEMIDESMOSE AND ASSOCIATED LATE-ONSET MUSCULAR DYSTROPHY.
CC -1- SIMILARITY: CONTRAINS 1 ACTIN-BINDING DOMAIN.
CC -1- SIMILARITY: CONTRAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
CC -1- SIMILARITY: CONTRAINS 3 PLECTIN REPEATS.
CC -1- SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z54367; CAA91196.1; -.
DR EMBL; U53204; AAB05427.1; -.
DR EMBL; U63610; AAB05428.1; -.
DR EMBL; U63609; AAB05428.1; JOINED.
DR EMBL; X97053; CAA65765.1; -.
DR HSSP; Q01082; 1BRK.
DR Genew; HGNC:9069; PLECL1.
DR MIM; 601282; -.
DR MIM; 226670; -.
DR InterPro; IPR001589; Actbind actinin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR001101; Plectin_repeat.
DR InterPro; IPR005326; S10 plectin.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00681; plectin; 19.
DR Pfam; PF03501; S10_plectin; 1.
DR SMART; SM00033; CH; 2.
DR SMART; SM00250; PLEC; 33.
DR SMART; SM00150; SPEC; 5.
DR PROSITE; PS00019; ACTININ_1; FALSE NEG.
DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
DR PROSITE; PS50021; CH; 2.
KW Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding;
KW Phosphorylation; Alternative splicing; Epidermolysis bullosa;
KW Disease mutation.
FT DOMAIN 1 1470 GLOBULAR 1.
FT DOMAIN 1471 2755 CENTRAL FIBROUS ROD DOMAIN.
FT DOMAIN 2756 4684 GLOBULAR 2.
FT DOMAIN 175 400 ACTIN-BINDING.
FT DOMAIN 179 282 CH 1.
FT DOMAIN 295 397 CH 2.
FT REPEAT 645 710 SPECTRIN 1.

```

```

FT REPEAT 740 824 SPECTRIN 2.
FT REPEAT 837 930 SPECTRIN 3.
FT REPEAT 1315 1415 SPECTRIN 4.
FT DOMAIN 1469 2756 COILED COIL (POTENTIAL).
FT REPEAT 2826 2863 PLECTIN 1.
FT REPEAT 2864 2901 PLECTIN 2.
FT REPEAT 2902 2939 PLECTIN 3.
FT REPEAT 2940 2977 PLECTIN 4.
FT REPEAT 2981 3015 PLECTIN 5.
FT REPEAT 3116 3153 PLECTIN 6.
FT REPEAT 3154 3191 PLECTIN 7.
FT REPEAT 3192 3229 PLECTIN 8.
FT REPEAT 3230 3267 PLECTIN 9.
FT REPEAT 3268 3305 PLECTIN 10.
FT REPEAT 3306 3343 PLECTIN 11.
FT REPEAT 3485 3522 PLECTIN 12.
FT REPEAT 3523 3560 PLECTIN 13.
FT REPEAT 3561 3598 PLECTIN 14.
FT REPEAT 3599 3636 PLECTIN 15.
FT REPEAT 3640 3674 PLECTIN 16.
FT REPEAT 3820 3857 PLECTIN 17.
FT REPEAT 3858 3895 PLECTIN 18.
FT REPEAT 3896 3933 PLECTIN 19.
FT REPEAT 3934 3971 PLECTIN 20.
FT REPEAT 3975 4008 PLECTIN 21.
FT REPEAT 4063 4100 PLECTIN 22.
FT REPEAT 4101 4138 PLECTIN 23.
FT REPEAT 4139 4176 PLECTIN 24.
FT REPEAT 4177 4214 PLECTIN 25.
FT REPEAT 4218 4252 PLECTIN 26.
FT REPEAT 4265 4305 PLECTIN 27.
FT REPEAT 4319 4356 PLECTIN 28.
FT REPEAT 4408 4445 PLECTIN 29.
FT REPEAT 4446 4483 PLECTIN 30.
FT REPEAT 4484 4521 PLECTIN 31.
FT REPEAT 4522 4559 PLECTIN 32.
FT REPEAT 4560 4597 PLECTIN 33.
FT DOMAIN 4250 4300 BINDING TO INTERMEDIATE FILAMENTS (BY SIMILARITY).
FT DOMAIN 4625 4640 4 X 4 AA TANDEN REPEATS OF G-S-R-X.
FT MOD_RES 4539 4539 PHOSPHORYLATION (BY CDC2) (BY SIMILARITY).
FT VARSPLIC 1 174 MVMAGMLMRDQLRAIYELFPREGVAVAKDORPRSLPHV
FT GVTNLQWBRAMASLRARGLVAFETPAMGCFPMYLTNEGIAHL
FT ROYLHPEIYASLQRRVRVAVMARRPRPHVQAQGPRL
FT GSPPKGGPLTEBQRLTKRELEVSPEITVPATQETTLA
FT RPEGPAPAT -> MSGEDAEVAVSDEVSNGSGSPSPD
FT TLPNLTGQRRSRSGGAGSGSVLPDAERAVIRIA (IN
FT ISOFORM 2 AND ISOFORM 3).
FT MISSING (IN ISOFORM 3).
FT L -> LL (IN MD-EBS).
FT /FTID=VAR_011336.
FT VARIANT 409 412 MISSING (IN MD-EBS).
FT VARIANT 429 429 /FTID=VAR_011337.
FT VARIANT 1003 1005 MISSING (IN MD-EBS).
FT CONFLICT 185 185 S -> F (IN REF. 2).
FT CONFLICT 259 259 N -> D (IN REF. 2).
FT CONFLICT 550 550 H -> N (IN REF. 2).
FT CONFLICT 560 560 I -> V (IN REF. 2).
FT CONFLICT 706 706 Q -> R (IN REF. 2).
FT CONFLICT 886 886 N -> Y (IN REF. 2).
FT CONFLICT 1002 1002 V -> A (IN REF. 2).
FT CONFLICT 1309 1309 V -> L (IN REF. 2).
FT CONFLICT 1321 1321 L -> V (IN REF. 2).
FT CONFLICT 1334 1334 V -> L (IN REF. 2).
FT CONFLICT 1534 1534 I -> M (IN REF. 2).
FT CONFLICT 1662 1662 A -> T (IN REF. 2).
FT CONFLICT 1688 1690 WLC -> RLR (IN REF. 2).
FT CONFLICT 1767 1767 Q -> E (IN REF. 2).
FT CONFLICT 1789 1789 L -> A (IN REF. 2).
FT CONFLICT 1910 1910 K -> R (IN REF. 2).
FT CONFLICT 2154 2154 K -> N (IN REF. 2).
FT CONFLICT 2160 2160 S -> R (IN REF. 2).
FT CONFLICT 2215 2215 R -> Q (IN REF. 2).

```

```

FT CONFLICT 2244 2244 S -> A (IN REF. 2).
FT CONFLICT 3027 3027 K -> E (IN REF. 2).
FT CONFLICT 3310 3310 E -> A (IN REF. 2).
FT CONFLICT 3361 3361 F -> L (IN REF. 2).
FT CONFLICT 3408 3408 F -> L (IN REF. 2).
FT CONFLICT 3447 3447 S -> A (IN REF. 2).
FT CONFLICT 3531 3531 G -> A (IN REF. 2).
FT CONFLICT 3580 3580 R -> S (IN REF. 2).
FT CONFLICT 3589 3589 K -> Q (IN REF. 2).
FT CONFLICT 3596 3596 E -> Q (IN REF. 2).
FT CONFLICT 3616 3616 N -> H (IN REF. 2).
FT CONFLICT 3686 3686 V -> A (IN REF. 2).

Query Match 5.3%; Score 106.5; DB 1; Length 4684;
Best Local Similarity 23.5%; Pred. No. 50;
Matches 70; Conservative 52; Mismatches 119; Indels 57; Gaps 13;

Qy 122 QLHPEA-IAAIOKALPFSKAEQILSKVSTGPTLETFDOLHRHPDAFYLAATAKALQ 180
Db 2598 KLQGEAKLQKSEBMQTVQGEQLLEFTQALQSFLEKSDILQR--ERFIEQKAKLEQ 2655
Qy 181 AHM-QLMKQVYLLDDQYVQPLPKGDQVLFNSDAEDLIDSKLKMREDEVLEHLMVADR 239
Db 2656 LFQDEVAKAQLREBQQRQ-----QQMEQERQRLV--ASMEBARRR--QHEAEQGVRR 2705
Qy 240 OKREIROLQELHKKOVLVDSITGMSPPDNO---TLAVLRGWRVRLRSREITLGR 295
Db 2706 KQEBQQLQEQRRQOEELAE-----ENQRLREQQLLEBQCHRAALAHSEVTASQ 2756
Qy 296 --ATKQNDIVDLLESGPAWK-----ISRKQVYKLKNNGDFPIANEGRRPI---- 340
Db 2757 VAATKTLPLNGRD-ALDGPAAABEHSFDGLRRKVSQRLQEG-ILSAEELQRLAQGHT 2814
Qy 341 -----YIDRPVLCGSKRLSNNSVVEISLRVFLINDLILIRAEAA 385
Db 2815 TVDEIARDEVRYHLQGRSSYIGLLKATNEKLSVYALQRLSPGATILIRQAA 2872

RESULT 3
DMD_HUMAN STANDARD; PRT; 3685 AA.
ID DMD_HUMAN
AC P11532; Q14169; Q14170;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dystrophin.
GN DMD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88194521; PubMed=3282674;
RA Koenig M., Monaco A.P., Kunkel L.M.;
RT "The complete sequence of dystrophin predicts a rod-shaped
RT cyoskeletal protein.";
RL Cell 53:219-228(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345106; PubMed=2668885;
RA Rosenthal A., Speer A., Billowitz H., Cross G.S., Forrest S.N.,
RA Davies K.E.;
RT "Two human cDNA molecules coding for the Duchenne muscular dystrophy
RT (DMD) locus are highly homologous.";
RL Nucleic Acids Res. 17:5391-5391(1989).
RN [3]
RP SEQUENCE OF 1-497 FROM N.A.
RX MEDLINE=87273512; PubMed=3607877;
RA Koenig M., Hoffman E.P., Bertelson C.J., Monaco A.P., Feener C.,
RA Kunkel L.M.;
RT "Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
RT preliminary genomic organization of the DMD gene in normal and

```

RT affected individuals."

RL Cell 50:509-517(1987).

RN [4]

RP SEQUENCE OF 404-1137 FROM N.A.

RX MEDLINE=8811512; PubMed=3428261;

RA Cross G.S., Speer A., Rosenthal A., Forrest S.M., Smith T.J.,

RA Edwards Y., Flint T., Hill D., Davies K.E.;

RT "Deletions of fetal and adult muscle cDNA in Duchenne and Becker

RT muscular dystrophy patients.";

RL EMO J. 6:3277-3283(1987).

RN [5]

RP SEQUENCE OF 665-722: 2098-2204 AND 2305-2366 FROM N.A.

RX MEDLINE=89083552; PubMed=3305741;

RA Chamberlain J.S., Gibbs R.A., Ranier J.A., Nguyen P.N.,

RA Caskey C.T.;

RT "Deletion screening of the Duchenne muscular dystrophy locus via

RT multiplex DNA amplification.";

RL Nucleic Acids Res. 16:1141-1156(1988).

RN [6]

RP SEQUENCE OF 2147-2204 FROM N.A.

RX MEDLINE=89345155; PubMed=2569720;

RA Blonden L.A.J., den Dunnen J.T., van Paassen H.M.B.,

RA Wagenaar M.C., Goetscholtzen P.M., Ginjaar H.B., Bakker E.,

RA Pearson P.L., van Ommen G.J.B.;

RT "High resolution deletion breakpoint mapping in the DMD gene by whole

RT cosmid hybridization.";

RL Nucleic Acids Res. 17:5611-5621(1989).

RN [7]

RP SEQUENCE OF 2305-2366 FROM N.A.

RA Huch A., Will K., Speer A., Bauer D.;

RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.

RN [8]

RP REVIEW ON DMD POINT MUTATION VARIANTS.

RX MEDLINE=95038763; PubMed=7951253;

RA Roberts R.G., Gardner R.J., Bobrow M.;

RT "Searching for the 1 in 2,400,000: a review of dystrophin gene point

RT mutations.";

RL Hum. Mutat. 4:1-11(1994).

RN [9]

RP REVIEW ON VARIANTS.

RX MEDLINE=94320940; PubMed=8045556;

RA Rittinsand F., Reiss J.;

RT "Microdeletions and polymorphisms in the Duchenne/Becker muscular

RT dystrophy gene.";

RL Hum. Genet. 94:111-116(1994).

RN [10]

RP VARIANT DMD ARG-54.

RX MEDLINE=94004962; PubMed=8401582;

RA Prior T.W., Papp A.C., Snyder P.J., Burghes A.H.M., Bartolo C.,

RA Sedra M.S., Western L.M., Mendell J.R.;

RT "A missense mutation in the dystrophin gene in a Duchenne muscular

RT dystrophy patient.";

RL Nat. Genet. 4:357-360(1993).

RN [11]

RP VARIANTS DMD HIS-365, TRP-2191 AND ARG-2937.

RX MEDLINE=95152525; PubMed=7849724;

RA Nigro V., Nigro G., Esposito M.G., Comi L.I., Molinari A.M.,

RA Pucca G.A., Pollitano L.;

RT "Novel small mutations along the DMD/BMD gene associated with

RT different phenotypes.";

RL Hum. Mol. Genet. 3:1907-1908(1994).

RN [12]

RP ALTERNATIVE SPLICING (DYSTROPHIN-1 AND -2).

RC TISSUE=Retina;

RA White R.A.;

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

RN [13]

RP ALTERNATIVE SPLICING (DYSTROPHIN-3).

RC TISSUE=Brain;

RX MEDLINE=89181947; PubMed=2648158;

RA Feener C.A., Koenig M., Kunkel L.M.;

RT "Alternative splicing of human dystrophin mRNA generates isoforms at

RT the carboxy terminus.";

RL Nature 338:509-511(1989).

CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE

CC PLASMA MEMBRANE.

CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS ARE PRODUCED BY

CC ALTERNATIVE SPLICING; THE 3 SHORTER VARIANTS ARE CALLED

CC DYSTROPHIN-1, -2 AND -3.

CC -1- DISEASE: Defects in DMD are the cause of Duchenne muscular

CC dystrophy (DMD) and Becker muscular dystrophy (BMD). DMD is the

CC most common form of muscular dystrophy; a sex-linked recessive

CC disorder. It typically presents in boys aged 3 to 7 year as

CC proximal muscle weakness causing waddling gait, toe-walking,

CC lordosis, frequent falls, and difficulty in standing up and

CC climbing up stairs. The pelvic girdle is affected first, then the

CC shoulder girdle. Progression is steady and most patients are

CC confined to a wheelchair by age of 10 or 12. Flexion contractures

CC and scoliosis ultimately occur. About 50% of patients have a lower

CC IQ than their genetic expectations would suggest. There is no

CC treatment. BMD resembles DMD in hereditary and clinical features

CC but is later in onset and more benign.

CC -1- DISEASE: Defects in DMD are a cause of X-linked dilated

CC cardiomyopathy (XLCM).

CC -1- MISCELLANEOUS: THE DMD GENE IS THE LARGEST KNOWN GENE IN HUMANS.

CC IT IS 2.4 MILLION BASE-PAIRS IN SIZE AND COMPRISES 79 EXONS.

CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY

CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,

CC ABR-120, ABR-180, OR BETA-FODRIN).

CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.

CC -1- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.

CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.

CC -1- DATABASE: NAME=DMD; NOTE=Dystrophin Mutation Database;

CC WWW="http://www.dmd.nl/database.html".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: X06179; CAA29545.1; ALT\_SEQ.

CC EMBL: X06178; CAA29544.1; -.

CC EMBL: X14298; CAA32479.1; ALT\_SEQ.

CC EMBL: X15495; CAA33518.1; -.

CC EMBL: X54820; CAA38589.1; -.

CC EMBL: X13045; CAA31451.1; -.

CC EMBL: X13046; CAA31452.1; -.

CC EMBL: X13047; CAA31453.1; -.

CC EMBL: X13048; CAA31454.1; -.

CC EMBL: U27203; AA86115.1; -.

CC EMBL: U27203; AA86116.1; -.

CC EMBL: X15148; CAA33245.1; -.

CC PIR: A27605; A27605.

CC PIR: A27162; A27162.

CC PIR: S05281; S05281.

CC HSPF: P46939; IQAG.

CC GeneW: HGNC:2928; DMD.

CC MIM: 300377; -.

CC MIM: 310200; -.

CC MIM: 300376; -.

CC MIM: 302045; -.

CC InterPro: IPR001589; Actbind\_actin.

CC InterPro: IPR001715; Calponin-like.

CC InterPro: IPR002017; Spectrin.

CC InterPro: IPR002349; WW.

CC InterPro: IPR001202; WW\_Rsp5\_WWP.

CC InterPro: IPR000433; ZnF\_ZZ.

CC Pfam: PF00307; CH; 2.

CC Pfam: PF00397; WW; 1.

CC Pfam: PF00435; Spectrin; 22.

CC Pfam: PF00569; ZF; 1.

CC PRINTS: PR00403; WWDOMAIN.





```

FT REPEAT 1934 2044 SPECTRIN 10.
FT REPEAT 2262 2282 SPECTRIN 11.
FT REPEAT 2276 2297 SPECTRIN 12.
FT REPEAT 2400 2509 SPECTRIN 13.
FT REPEAT 2557 2617 SPECTRIN 14.
FT REPEAT 2620 2727 SPECTRIN 15.
FT REPEAT 2730 2837 SPECTRIN 16.
FT REPEAT 2840 2944 SPECTRIN 17.
FT REPEAT 2946 2972 SPECTRIN 18.
FT REPEAT 2986 3029 SPECTRIN 19.
FT REPEAT 3086 3162 SPECTRIN 20.
FT REPEAT 3195 3273 SPECTRIN 21.
FT REPEAT 3276 3382 SPECTRIN 22.
FT REPEAT 3385 3491 SPECTRIN 23.
FT REPEAT 3494 3518 SPECTRIN 24.
FT REPEAT 3603 3709 SPECTRIN 25.
FT REPEAT 3722 3817 SPECTRIN 26.
FT REPEAT 3869 3930 SPECTRIN 27.
FT REPEAT 3933 4039 SPECTRIN 28.
FT REPEAT 4042 4149 SPECTRIN 29.
FT REPEAT 4152 4258 SPECTRIN 30.
FT REPEAT 4261 4368 SPECTRIN 31.
FT REPEAT 4371 4477 SPECTRIN 32.
FT REPEAT 4480 4587 SPECTRIN 33.
FT REPEAT 4594 4695 SPECTRIN 34.
FT REPEAT 4698 4804 SPECTRIN 35.
FT REPEAT 4807 4872 SPECTRIN 36.
FT REPEAT 4917 4941 SPECTRIN 37.
FT CA_BIND 4987 4998 EF-HAND 1 (POTENTIAL).
FT CA_BIND 5023 5034 EF-HAND 2 (POTENTIAL).
FT DOMAIN 5173 5180 POLY-SER.
FT DOMAIN 5252 5267 4 X 4 AA TANDEN REPEATS OF [GS]-S-R-[AR].
FT VARSPLIC 1 73 MSSSEETLSESRCSRSRCSRSRSGSLSPCPD
FT VARSPLIC 1 181 TLPMLPLHEQKRSQSDVLDPAERAVRVA -> EKEFV
FT VARSPLIC 1 181 AAEVDLEERYK (IN ISOFORM 1).
FT REPEAT 2286 2306 MSSSEETLSESRCSRSRCSRSRSGSLSPCPD
FT REPEAT 2310 2330 TLPMLPLHEQKRSQSDVLDPAERAVRVADEEDRQCK
FT REPEAT 2334 2354 TTPKAVNGLMKRKHNDLYEDLDKQHVNLISLLEYSGIK
FT REPEAT 2358 2378 LPREKGRMFRHLQNVQIALDFLQKQOYVLVIRNDITDGE
FT REPEAT 2382 2402 NPKTLGLIWLITLHFQ -> MGNISGCVKEPKESIAVBEK
FT REPEAT 2406 2426 APISPEKRVRFGRKRGKILTPPEASHHEALEGTVLEET
FT REPEAT 2430 2450 ETLTLTLARLPEKPGVGAHEHPDILPFGDAPNSGVGD
FT REPEAT 2454 2474 GMIVQKSFQAEIQTALHLENSSVGGAGDSLEGGTV
FT REPEAT 2478 2498 IAHLDNPERNCEKSVQLEFPPTACSSRAVILLPLDGE
FT REPEAT 2502 2522 TAVEQGTLLRHRHRSSTLPTDYSEVTLVDQOPEBGSVG
FT REPEAT 2526 2546 GRTKSVSPAPPTGWSIACSVASSIPKSGDPIHTEPTHVG
FT REPEAT 2550 2570 LVSCKPIPNASQDSLVSIGTIVSLIPSSGYSGLRLHG
FT REPEAT 2574 2594 IRPEDTEPKTSTPSEEDGTLSL (IN ISOFORM 3).
FT CONFLICT 393 393 L -> P (IN REF. 2; AAC5290).
FT CONFLICT 1882 1882 Q -> H (IN REF. 2).
FT SEQUENCE 5327 AA; 607972 MW; 5DEAF56514BFPA CRC64;

Query Match 5.2%; Score 103.5; DB 1; Length 5327;
Best Local Similarity 21.6%; Pred. No. 97;
Matches 71; Conservative 59; Mismatches 124; Indels 75; Gaps 15

QY 52 QPLQYTKDLGRWKRPANDLLINAVIQTNDLTLSVHLGVKFSCTFLREYQ-----ERMVYA 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 597 EEMQMKLEBAEN--GNDLPSVELQLETOQ--NHTSVSE-ELSSSVKEARLYEKGSKQNPH 651

QY 106 LIYDVIYSKLAQ-----AMRQLHPAIALAIGSKA-----LFSKABEOLISKVSGT 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 652 TSIVYETLIGLKYQCKLKETSTSFMRHLOSLHKFVSRAAEILWLNGKEEELIACD--WSD 710

QY 152 SQPT-----LETFODL-----LHRHPDAFYLTARTAKLOAHWQL 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 711 SNPNISAKKTVSELTMELEEGKODVFRSLQDTRAEVILSTNNHPKQOTVEAYSAVQSOLOM 770

QY 186 MKQYVLLDEQTVQPLPKGDQVNLFSBAEDLIDSKDKMDREVLLEHILVAVDRQ--KRE 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 771 MKQDLCLCVQHQKEMAAVYQF--FSFARLD--ESFLRNLDQSL--KRXXTATADRSTLSL 824

QY 244 IRQLEQELHKWVLVDSTIGMSSPDFDNQTLAVLRGRMYRYLMRSF-----EI 291

```

D6	825	EDLQDSNDEKEHQGLIQSSSVASLVGRSKTIVQLKPRNPDHVLKSTLSVKAI CDYRQLEI	884
Oy	292	TLGATKDNOQIDVDLSLEGPAWKISRKOG 320	
		:   : : :   :	
Db	885	TI---CKNDECVELDNSORTKMKVISPFG 910	
 RESULT 5			
GfRA	BORBU	STANDARD;	PRT; 810 AA.
ID	_GYRA_BORBU	051396; Q44931;	
AC		15-DEC-1998 (Rel. 37, Created)	
DT		15-DEC-1998 (Rel. 37, Last sequence update)	
DT		16-OCT-2001 (Rel. 40, Last annotation update)	
DE		DNA gyrase subunit A (EC 5.99.1.3).	
GN		GYRA OR BB0435.	
OS		Borrelia burgdorferi ( Lyme disease spirochete ).	
OC		Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.	
OX		NCBI_TaxId=139;	
RN	[1]	SEQUENCE FROM N.A.	
RP		STRAIN=ATCC 35210 / B31;	
RC		MEDLINE=98065943; PubMed=9403685;	
RX		Fraser C.M., Chaiens S., Huang W.M., Sutton G.G., Clayton R.A.,	
RA		Lathiera R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwin M.,	
RA		Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,	
RA		Petereson J., Keilavage A.R., Quackenbush J., Salzberg S., Hanson M.,	
RA		van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,	
RA		Uterback T., Matthey L., McDonald L., Artiach P., Bowman C.,	
RA		Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,	
RA		Smith H.O., Venter J.C.;	
RT		"Genomic sequence of a Lyme disease spirochaete, Borrelia	
RL		burgdorferi.";	
RL	Nature	390:580-586(1997).	
RN	[2]	SEQUENCE OF 450-545 FROM N.A.	
RP		STRAIN=212;	
RC		MEDLINE=93146383; PubMed=1490605;	
RX		Old I.G., Macdougall J., Saint-Girons I., Davidson B.E.;	
RT		"Mapping of genes on the linear chromosome of the bacterium Borrelia	
RL		burgdorferi: possible locations for its origin of replication.";	
FEMS	Microbiol. Lett.	78:245-250(1992).	
-1-	FUNCTION:	DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCUAR DOUBLE-	
CC		STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE	
CC		INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED	
CC		DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.	
CC	-1-	CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining	
CC		of double-stranded DNA.	
CC	-1-	SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA	
CC		BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE	
CC		ENZYME FORMS AN A2B2 TETRAMER.	
CC		-----	
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC		the European Bioinformatics Institute. There are no restrictions on its	
CC		use by non-profit institutions as long as its content is in no way	
CC		modified and this statement is not removed. Usage by and for commercial	
CC		entities requires a license@isb-sib.ch (See http://www.isb-sib.ch/announce/	
CC		cc send an email to license@isb-sib.ch).	
CC		-----	
CC	EMBL; AB001148; AAC66803.1; --		
DR	EMBL; Z12165; CAAT78157.1; --		
DR	HSSP; P09097; IAB4.		
DR	TIGR; BB0435; --		
DR	InterPro: IPR002205; DNA_topoisotv.		
DR	pfam; PF00521; DNA_topoisotv. 1.		
DR	PfDom; PD000742; DNA_topoisotv; 1.		
DR	SMART; SM00434; TOP4C; 1.		
DR	TIGRFAMs; TIGR01063; GYRA. 1.		
DR	Topoisomerase; Isomerase; ATP-binding; Antibiotic resistance;		
KW	Complete proteome.		
FT	ACT_SITE 124 124	DNA_CLEAVAGE (BY SIMILARITY) .	

```

FT CONFLICT 469 472 IREE -> YKER (IN REF. 2).
PT CONFLICT 539 539 F -> Y (IN REF. 2).
SQ SEQUENCE 810 AA; 91379 MW; 395F3930F0EF1650 CRC64;

Query Match 5.2%; Score 103; DB 1; Length 810;
Best Local Similarity 21.0%; Pred. No. 7.9;
Matches 81; Conservative 65; Mismatches 131; Indels 108; Gaps 18;

QY 67 NDLLINAVLDTNDTSVHLS-----VKRSCFTREVERKVALYLPVSKLACQ 118
DB 334 NMLALVNCIPQOLNEELLFEIERKNIIRRIEFDLKAKAKA-----VLEGNI- 386
QY 119 AMRQLHPEAIAIOSKALFSAKEBOLSKVSTSQPLETFODLHHPDAFYLAATAKA 178
DB 387 ALNNI-DEVITIKISSKLAKARELVSNFG-LSIETQANSVLDMLQKLTLEIFKLEE 444
QY 179 LOAHMQLKQYLLLEDQVQPLPKDQVYN-----FSD--AEDLIDSKLD 223
DB 445 LNILSLTKDY---EDILINPV---RIINIRIETINLGLKFGDERRTKIYDEEVLKT 497
QY 224 MRDEVLHEHLVADROKREIRLOLEQLHKQVLDVSTGMSPPDQTLAVL----- 277
DB 498 SMSDMQKENVMTTKGFLKLSQNEKQDGTGCK--GLSSFDLNDGDEIVIALCVNT 555
QY 278 -----RGRMYRYLNRSE-----ITLG-----RATKDNQIDV 304
DB 556 HDYLEFMISNEGL--YLINAYIKSSRSASKQNSIELINDGDEILITKSKDITDPA 613
QY 305 DLSLEGPAKISR-----KQGVITLK--NNGDFPIANEGRPPIYIDGRPVLCGSK- 352
DB 614 YLLTTAGSKLARFEFSTDFKAVKSRGVIVIKLNDKDFVTSAE---IVFKDEKVICLSKK 669
QY 353 -----WELSNVVEIASLR 367
DB 670 GSAPFENSRDVLNTRGTQVCGMK 694

RESULT 6
DMD_MOUSE STANDARD; PRT; 3678 AA.
ID_DMD_MOUSE
AC P1531; G60703;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dystrophin.
GN DMD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92253376; PubMed=1579466;
RA Bies R.D., Phelps S.F., Cortez M.D., Roberts R., Caskey C.T.,
RA Chamberlain J.S.;
RT "Human and murine dystrophin mRNA transcripts are differentially
RT expressed during skeletal muscle, heart, and brain development.";
RL Nucleic Acids Res. 20:1725-1731 (1992).
RN [2]
RP SEQUENCE OF 1-201 FROM N.A.
RX MEDLINE=87273512; PubMed=3607877;
RA Koenig M., Hoffman E.P., Bertelson C.J., Monaco A.P., Feener C.,
RA Kunzel L.M.;
RT "Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
RT preliminary genomic organization of the DMD gene in normal and
RT affected individuals.";
RL Cell 50:509-517 (1987).
RN [3]
RP SEQUENCE OF 300-1390 FROM N.A.
RX MEDLINE=88018015; PubMed=3659917;
RA Hoffman E.P., Monaco A.P., Feener C.C., Kunzel L.M.;
RT "Conservation of the Duchenne muscular dystrophy gene in mice and
RT humans.";

```

```

RL Science 238:347-350 (1987).
RN [4]
RP SEQUENCE OF 986-1056 FROM N.A.
RC STRAIN=C57BL/10; TISSUE=Skeletal muscle;
RX MEDLINE=94154933; PubMed=811539;
RA Chamberlain J.S., Phelps S.F., Cox G.A., Maichele A.J.,
RA Greenwood A.D.;
RT "PCR analysis of muscular dystrophy in mdx mice.";
RL Mol. Cell Biol. Hum. Dis. Ser. 3:167-169 (1993).
RN [5]
RP ALTERNATIVE SPLICING.
RC STRAIN=C57BL/10; TISSUE=Retina;
RX MEDLINE=95360002; PubMed=7633443;
RA D'Souza V.N., Nguyen T.M., Morris G.E., Karges W., Pillers D.-A.M.,
RA Ray P.N.;
RT "A novel dystrophin isoform is required for normal retinal
RT electrophysiology.";
RL Hum. Genet. 4:837-842 (1995).
CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
CC PLASMA MEMBRANE.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 11 DISTINCT ISOFORMS EXIST.
CC -1- TISSUE SPECIFICITY: DIFFERENTIALLY EXPRESSED DURING SKELETAL
CC MUSCLE, HEART, AND BRAIN DEVELOPMENT. ALSO EXPRESSED IN
CC RETINA.
CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
CC ABP-120, ABP-180, OR BETA-FODRIN).
CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
CC -1- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.
CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M68859; AAB02797.1; -
CC EMBL; M18025; AAA37530.1; -
CC EMBL; U56724; AAB01216.1; -
CC EMBL; U15218; AAA87068.1; -
CC PIR; B27162; B27162.
CC HSSP; P46939; IQAG.
CC MGD; MGI:94909; Dmd.
CC InterPro: IPR001589; Actbind actnin.
CC InterPro: IPR001715; Calponin-like.
CC InterPro: IPR002017; Spectrin.
CC InterPro: IPR002349; WW_Reps_WWP.
CC InterPro: IPR001202; WW_Domain_2.
CC Pfam; PF00307; CH; 2.
CC Pfam; PF00397; WW; 1.
CC Pfam; PF00435; spectrin; 21.
CC Pfam; PF00569; Z2; 1.
CC PRINTS; PR00403; WWDOMAIN.
CC SMART; SM00033; CH; 2.
CC SMART; SM00150; SPEC; 20.
CC SMART; SM00291; Znf_Z2; 1.
CC PROSITE; PS00019; ACTININ_1; 1.
CC PROSITE; PS00020; ACTININ_2; 1.
CC PROSITE; PS01159; WW_Domain_1; 1.
CC PROSITE; PS50020; WW_Domain_2; 1.
CC PROSITE; PS01357; ZF_Z2_1; 1.
CC PROSITE; PS01357; ZF_Z2_2; 1.
CC Structural Protein; Actin-binding; Calcium-binding; Cytoskeleton;
CC Repeat; Alternative splicing; ACTIN-BINDING.
FT DOMAIN 1 240
FT DOMAIN 15 119 CH 1.

```

```

FT DOMAIN 134 237 CH 2.
FT REPEAT 341 449 SPECTRIN 1.
FT REPEAT 450 558 SPECTRIN 2.
FT REPEAT 561 669 SPECTRIN 3.
FT REPEAT 721 830 SPECTRIN 4.
FT REPEAT 832 936 SPECTRIN 5.
FT REPEAT 945 1047 SPECTRIN 6.
FT REPEAT 1050 1156 SPECTRIN 7.
FT REPEAT 1159 1265 SPECTRIN 8.
FT REPEAT 1268 1369 SPECTRIN 9.
FT REPEAT 1470 1570 SPECTRIN 10.
FT REPEAT 1573 1678 SPECTRIN 11.
FT REPEAT 1681 1782 SPECTRIN 12.
FT REPEAT 1879 1981 SPECTRIN 13.
FT REPEAT 2013 2103 SPECTRIN 14.
FT REPEAT 2106 2210 SPECTRIN 15.
FT REPEAT 2213 2318 SPECTRIN 16.
FT REPEAT 2468 2570 SPECTRIN 17.
FT REPEAT 2573 2679 SPECTRIN 18.
FT REPEAT 2682 2795 SPECTRIN 19.
FT REPEAT 2900 2900 SPECTRIN 20.
FT REPEAT 2902 2924 SPECTRIN 21.
FT REPEAT 2927 3033 SPECTRIN 22.
FT DOMAIN 3048 3081 MW.
FT ZN FING 3300 3347 ZZ-TYPE.
FT CONFLICT 463 D -> H (IN REF. 3).
FT CONFLICT 677 S -> F (IN REF. 1).
FT CONFLICT 2337 V -> L (IN REF. 1; AAB02797).
SQ SEQUENCE 3678 AA; 425810 MW; 1D2E74CFD8035EE CRC64;

Query Match 5.1%; Score 102; DB 1; Length 3678;
Best Local Similarity 20.9%; Pred. No. 75;
Matches 69; Conservative 60; Mismatches 121; Indels 80; Gaps 14;

QY 32 PVPSPAPAGLTTRVYKSKOPLOVTKDGRMKPANDLLINAVLOTNDLVHLGVKES 91
DB 3410 PVDASAPASPGSHDPTDRHIEHYASR-LAEMENSGSLINDISPNESIDDEHLILQHY 3468
QY 92 CRFTLEVERQWYALIDYVISTKLACQAMKQHPALAIQSAALFSAKEBOLSKVSGT 151
DB 3469 CO-----SLNQDPSLSQ-----PRSPAQILSLSEB 3494
QY 152 SQPTLE-TFQDILLHRPDAFYLAARTAKALQAHQMLKQYVLLDQTVQPLPKDQVLYNS 210
DB 3495 EKEGLERITLADLEEN-----RNLOAEYDRKKQOH-EHKGLSLPSPPEMPTIS 3542
QY 211 -----DAEDLIDDSKLKMDREVLHEHLMVADRQKREIROLEOELHAKQVIVD----- 259
DB 3543 PQSPDAE-LIAEAKLLRQHKGRLFARMQILBDHNK-----QLESQHLRLRLLEQPGQAA 3597
QY 260 --STGMSGSPDPDNOQLAVLRGAMVRYLMKRSREITLGRATKDNQIDVDLSLEGPAKISR 317
DB 3598 KVGNTVSSPSTSLORSDDSQPMLR-----VVGQSISEMGEDL-LSPPODTSTG 3648
QY 318 KQGVTKLKNNGDFEIANEGRRPRYIDGRPV 347
DB 3649 LEEVMEQLNNS--FSSSRGN---APGKEW 3673

POL_CAEVC STANDARD; PRT; 1109 AA.
ID POL_CAEVC
AC P33459;
DT 01-FEB-1994 (Rel. 28; Created)
DT 01-FEB-1994 (Rel. 28; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE POL. polyprotein (Contains: Protease (Retroposin) (EC 3.4.23.-);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.1.26.4)].
GN POL.
OS Caprine arthritis encephalitis virus (strain Cork) (CAEV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11661;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=91021037; PubMed=2171210;
RA Saitarelli M., Querat G., Konings D.A.M., Vigne R., Clements J.E.;
RT "Nucleotide sequence and transcriptional analysis of molecular clones
of CAEV which generate infectious virus.";
RL Virology 179:347-364(1990).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ (DNA)(N).
CC -1- P1M: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS PROTEIN MAY BE SYNTHESIZED AS A GAG-POL
POLYPROTEIN.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sdb.ch/announce/
or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: M33677; AAA91826.1; ALT_INIT.
CC PIR: B45345; B45345.
CC HSR: P16088; 1DUT.
CC MEROPS: A02.006; -.
DR InterPro: IPR001995; Asparticase_rtyv.
DR InterPro: IPR001969; Asparticase_site.
DR InterPro: IPR001428; DeoxyUTPase.
DR InterPro: IPR001037; Integrase_C.
DR InterPro: IPR003308; Integrase_Zn.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR00477; RYase.
DR InterPro: IPR001584; Rye.
DR Pfam: PF00075; rnaseH; 1.
DR Pfam: PF00078; rtyv; 1.
DR Pfam: PF00552; Integrase; 1.
DR Pfam: PF00665; rye; 1.
DR Pfam: PF00692; dUTPase; 1.
DR Pfam: PF02022; Integrase_Zn; 1.
DR ProDom: PD000946; deoxyUTPase; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS00175; ASP_PROT_RETROV; 1.
KW Hydrolase; Aspartyl protease; RNA-directed DNA polymerase;
KW Transferase; Nuclease; Endonuclease; Polypeptide.
FT CHAIN 1 152 PROTEINASE.
FT CHAIN 153 865 REVERSE TRANSCRIPTASE.
FT CHAIN 866 1109 ENDONUCLEASE.
FT ACT_SITE 68 68 BY SIMILARITY.
SQ SEQUENCE 1109 AA; 127678 MW; 97B2F4B370B03CF3 CRC64;

Query Match 5.1%; Score 101.5; DB 1; Length 1109;
Best Local Similarity 19.7%; Pred. No. 16;
Matches 81; Conservative 54; Mismatches 142; Indels 135; Gaps 20;

QY 21 EKKVSKAP-----STVPSPAPAGLTTRVYKSKOPLOVTKDGRMKPANDLLINAV 75
DB 190 ECKLKGKAPPHWTCTP-----FCIKKSGKMWMLIDFRELNK- 228
QY 76 LOTNDLTSVHL-----GVKFSCTFLAEVOERWAL-LYDVVISIKLACQAMROLHPAI 128
DB 229 -QTEBLTEAQQLPHPGGIQRKQHVITLIDIGAVFTIPVEP-YREYTCFTL-LSPNNL 284
QY 129 AAIOSKALFSAKEBOLSKVSGTSPQTL-ETFQDILLHRPD----- 168
DB 285 G--PCKRYWKLPQGWKLSPEVYQFTWGEILBWDVIGQHPLEQFGIYMDIYIGSLBETK 342
QY 169 -----AFYLA-----RTAKALQAHW---QIMKQYVLLDQTVQPLPKDQVYL 207
DB 343 KREIVKDLANYIAQYGTFLPEEKROKGYPAKMLGFELHPQTWKFQKTLPELTGKTITL 402

```

QY 208 NFSADLIDSKLMDVELEHNE-----LMVADRRQREIROLELHMKQ 255  
 DB 403 N-----KIQKVLGELVWRSIQSGSIPIPLIKLMWEGDRHQSRKLEHVAHWE 451  
 QY 256 VLVDSTGMSPPDPNQTAVLRGMVRYLRMSREITLGRATKQNOIDVLSLE--GPW 313  
 DB 452 ACRKKLEME-----NYNKKDKVYGQALMGDKAIEIYVQEKGPDLW 495  
 QY 314 -----KIRKQGVILKKNK--GDFIANEGRR--IYIDRPVLCSSKWL 355  
 DB 496 VNVVHNKLSIPQGVIAQKLTQEVILRTGKLPIWLLPGKE-----EDWRL 543

RESULT 8  
 ID\_DMD\_CANFA STANDARD; PRT; 3680 AA.  
 AC 097592;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Dystrophin.  
 GN DMD.  
 OS Canis familiaris (dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Golden retriever;  
 RA Cavillie K.S., Mann C.J., Schatzberg S.J., Milton S.D.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE  
 CC PLASMA MEMBRANE.  
 CC -1 SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY  
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIBRININ,  
 CC APP-120, ABP-180, OR BETA-FODRIN).  
 CC -1 SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.  
 CC -1 SIMILARITY: CONTAINS 1 WW DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.  
 CC -1 SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>;  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----  
 CC EMBL; AF070485; AAC83646.1; --  
 DR HSSP; P46939; 10AG.  
 DR InterPro; IPR001589; Actbind actin.  
 DR InterPro; IPR001715; Calponin-like.  
 DR InterPro; IPR002017; Spectrin.  
 DR InterPro; IPR001202; WW Rep5\_WWP.  
 DR InterPro; IPR00433; Znf\_ZZ.  
 DR Pfam; PF00307; CH; 2.  
 DR Pfam; PF00397; WW; 1.  
 DR Pfam; PF00435; Spectrin; 21.  
 DR Pfam; PF00569; ZZ; 1.  
 DR SMART; SM00033; CH; 2.  
 DR SMART; SM00150; SPEC; 21.  
 DR SMART; SM00456; WW; 1.  
 DR SMART; SM00291; Znf\_ZZ; 1.  
 DR PROSITE; PS00019; ACTININ\_1; 1.  
 DR PROSITE; PS00020; ACTININ\_2; 1.  
 DR PROSITE; PS50021; CH; 2.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 1.  
 DR PROSITE; PS50020; WW\_DOMAIN\_2; 1.  
 DR PROSITE; PS01357; ZF\_ZZ\_1; 1.  
 DR PROSITE; PS50135; ZF\_ZZ\_2; 1.  
 KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;

KW Repeat; Zinc-finger.  
 FT DOMAIN 15 119 CH 1.  
 FT DOMAIN 134 237 CH 2.  
 FT REPEAT 134 237 SPECTRIN 1.  
 FT REPEAT 340 448 SPECTRIN 2.  
 FT REPEAT 449 557 SPECTRIN 3.  
 FT REPEAT 560 668 SPECTRIN 4.  
 FT REPEAT 720 829 SPECTRIN 5.  
 FT REPEAT 831 935 SPECTRIN 6.  
 FT REPEAT 944 1047 SPECTRIN 7.  
 FT REPEAT 1050 1156 SPECTRIN 8.  
 FT REPEAT 1159 1265 SPECTRIN 9.  
 FT REPEAT 1268 1369 SPECTRIN 10.  
 FT REPEAT 1470 1570 SPECTRIN 11.  
 FT REPEAT 1573 1678 SPECTRIN 12.  
 FT REPEAT 1681 1782 SPECTRIN 13.  
 FT REPEAT 1879 1981 SPECTRIN 14.  
 FT REPEAT 2013 2103 SPECTRIN 15.  
 FT REPEAT 2106 2210 SPECTRIN 16.  
 FT REPEAT 2213 2320 SPECTRIN 17.  
 FT REPEAT 2470 2572 SPECTRIN 18.  
 FT REPEAT 2575 2681 SPECTRIN 19.  
 FT REPEAT 2684 2797 SPECTRIN 20.  
 FT REPEAT 2800 2902 SPECTRIN 21.  
 FT REPEAT 2904 2926 SPECTRIN 22.  
 FT REPEAT 2929 3035 SPECTRIN 23.  
 FT DOMAIN 3050 3083 WW.  
 FT ZN\_FING 3302 3349 ZZ-TYPE.  
 SQ SEQUENCE 3680 AA; 425650 MW; 539F1C9D72377872 CRC64;

Query Match 5.1%; Score 101; DB 1; Length 3680;  
 Best Local Similarity 22.8%; Pred. No. 89;  
 Matches 66; Conservative 40; Mismatches 83; Indels 101; Gaps 15;

QY 16 EPSSEKKVSKASTPVPSPAPAGIT-----KRVKSKPLOTYTKLGRKK---P 65  
 DB 2399 EDLSMDWVTVTLQLELRKAKPGAPGLTVRAPPSQTVTLVTP--ATKETAISKLEMP 2457  
 QY 66 ANDLLINAVLQTN---DLTSVHLGVFSCFTLREVERVVALYDPVI----- 112  
 DB 2458 SSLLEVALDLDFRATWELTD-----WLSIL-DRAVKSQVWYGD 2497  
 QY 113 -----SKLCAQMRQLHPAIAIOLSKALFSKAEQLSKVSGTSQPLETFODL 162  
 DB 2498 LEDINEMITKQKATLQDLQGRPOLEBII-----TAQNLKNTKSQGEARTIT--DR 2548  
 QY 163 LHRHPDAFYLRKALQAHQWLMKQY-----LLEQOT--VQPLPKGDQYLANSD 211  
 DB 2549 IER-----IOSQWDEVOEHQNRRLQITETMLKDSITOWLEAKESAEQVLQGAR 2595

RESULT 9  
 SYS\_MYCGE STANDARD; PRT; 417 AA.  
 ID\_SYS\_MYCGE  
 AC P47251;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE SerV1-tRNA synthetase (EC 6.1.1.11) (Serine-tRNA ligase) (Serns).  
 GN SERN OR MG005.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NCBI\_TaxID=20997;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kierlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,

```

RA Nguyen D.T., Uteerback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Jancier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RA "The minimal gene complement of Mycoplasma genitalium.";
RA Science 270:397-403(1995).
RN [2]
RN RP SEQUENCE OF 1-125 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94364962; PubMed=8083173;
RA Bailey C.C., Bott K.F.;
RA "An unusual gene containing a dnaN N-terminal box flanks the putative
RT origin of replication of Mycoplasma genitalium.";
RL J. Bacteriol. 176:5814-5819(1994).
RN [3]
RN RP SEQUENCE OF 6-112 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RA "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate
CC + L-seryl-tRNA(Ser).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39679; AAC1221.1; -
DR EMBL: U09251; AA57073.1; -
DR EMBL: U02210; AAD12502.1; -
DR HSSP: P34945; 1SER.
DR TIGR: MG005; -
DR InterPro: IPR002106; AACRNA_ligaseII.
DR InterPro: IPR002314; tRNA-synt_2b.
DR InterPro: IPR002317; tRNA-synt_ser.
DR Pfam: PF00587; tRNA-synt_2b_1.
DR Pfam: PF02403; Seryl-tRNA_N_1.
DR PRINTS: PR00981; TRNASYNTHSER.
DR TIGRFAMs: TIGR00414; sers; 1.
DR PROSITE: PS50862; AA TRNA_LIGASE II; 1.
KW Aminocacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 417 AA; 47987 MW; 5E43280A5915DD0 CRC64;
Query Match 5.0%; Score 99.5; DB 1; Length 417;
Best Local Similarity 23.2%; Pred. No. 57;
Matches 41; Conservative 35; Mismatches 62; Indels 39; Gaps 8;
OY 215 LIDSKLQMDR-----EVLHEHLMVADRQKREIRFQELHKKQVLDVST 262
DB 1 MDPNKLNNVDFEKKLELRNVNEOLLNQIOTDKMKKNLQOLBLANQKOSILLAKOVA 60
OY 263 GMSPPFDNQTALVLRGVRVYLMRSREI-----TIGRATKQNO-IDVLSLEGPAWKIS 316
DB 61 KKK-----DKKKL-----LAESKELKQKLENNNNNAKOSNTSODLLNFP--NTA 104
OY 317 RKGQVILKNNGDFIFANEGRRPIYIDGRPV-----LCGSKMLSNNSVVEIASLRFV 369
DB 105 HESVPGKKNESANLELLKGRKRVF-DPKPLPHELCEKLNVAFAFKATKISTSTRV 160

```

```

DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-aminocyclopropane-2-carboxylate synthase 2 (EC 4.4.1.14) (ACC
DE synthase 2) (S-adenosyl-L-methionine methylthioadenosine-lyase 2).
GN ACS2 OR ACC1 OR AT1G01480 OR F2214.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eustoidia II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=93028584; PubMed=1357670;
RA van der Straeten D., Rodriguez-Pousada R.A., Villarroel R., Hanley S.,
RA Goodman H.M., van Montagu M.;
RA "Cloning, genetic mapping, and expression analysis of an Arabidopsis
RT thaliana gene that encodes 1-aminocyclopropane-1-carboxylate
RT synthase.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:9969-9973(1992).
RN [2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=93066381; PubMed=1438312;
RA Liang X.-W., Abel S., Keller J.A., Shen N.F., Theologis A.;
RA "The 1-aminocyclopropane-1-carboxylate synthase gene family of
RT Arabidopsis thaliana.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:11046-11050(1992).
RN [3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Terry N.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altabi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maltl R., Marziani A.,
RA Miltseher J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uteerback T., Van Aken S., Vayesberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
RN [5]
RN RP FUNCTION: CATALYZES THE FORMATION OF 1-AMINOCYCLOPROPANE-1-
RN CARBOXYLATE, A DIRECT PRECURSOR OF ETHYLENE IN HIGHER PLANTS.
RN -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = 1-
RN aminocyclopropane-1-carboxylate + methylthioadenosine.
RN -1- Cofactor: PYRIDOXAL PHOSPHATE.
RN -1- PATHWAY: Ethylene biosynthesis; first (rate limiting) step.
RN -1- SUBUNIT: HOMODIMER.
RN -1- TISSUE SPECIFICITY: HIGH IN DEVELOPING LEAVES AND IN FLOWERS.
RN -1- INDUCTION: BY ETHYLENE.
RN -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
RN AMINOTRANSFERASES.
RN -----
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC -----  
 DR EMBL: Z12614; CAA78260.1; -  
 DR EMBL: M95595; AAB59298.1; -  
 DR EMBL: M95594; AAB97516.1; -  
 DR EMBL: Y12776; CAA73310.1; -  
 DR EMBL: AC061957; AAF81308.1; -  
 DR PIR: S31646; S31646.  
 DR PIR: A47199; A47199.  
 DR PIR: S46190; S46190.  
 DR HSSP: P37821; 188G.  
 DR InterPro: IPR001176; ACC synthase.  
 DR InterPro: IPR004839; Amino transferfl/2.  
 DR InterPro: IPR004838; Nitransf.1.  
 DR Pfam: PF00155; aminotran.1.2; 1.  
 DR PRINTS: PR00753; ACCSYNTBASE.  
 DR PROSITE: PS00105; AA TRANSFER CLASS 1; 1.  
 KM Fruit ripening; Ethylene biosynthesis; lyase; Pyridoxal phosphate;  
 KM Multigene family.  
 FT BINDING 279  
 FT VARIANT 136 136 M->I.  
 SQ SEQUENCE 496 AA; 55531 MW; 766318A9B5F1566 CRC64;

Query March 5.0%; Score 99.5; DB 1; Length 496;  
 Best Local Similarity 20.2%; Pred. No. 7.2; Indels 95; Gaps 12;  
 Matches 68; Conservative 56; Mismatches 117;

Qy 32 PVPSPAPAPGLTK-----RVKSKQPLQVTKDGRWKPANDLLINAVLOTNDLTSL 86  
 Db 170 PVPSSSNFPLTVDAAEMAYKKAQSNKKYKGLLTNPSPNL---GMLDKDITLNL-- 224  
 Qy 87 GVKSKRFTLEVRQERWYLLY---DPVISKACQAMRQLHPKALIAIQSKALFSKAE 142  
 Db 225 -VFETRNKIHVVDEIYAATVFAAGDFV-----SVAEVVDVDSIEVNV 268  
 Qy 143 QLSKVSQSGP-----TLTFEODLL---HHPPAFYLAFAKXLAQHWOMKQY 190  
 Db 269 DLHIVYLSLQMDGLPGRVGVISFNDVSVSCAKMSSFGI-----VSSQTLMLASM 322  
 Qy 191 LLEDOVQPLPKGDVQVLFNSDAEDLIDSKLMDREVLLEHLMVAD----- 237  
 Db 323 LSDQDFV-----NF-----LMSSRLRGIRHKYFTTGIRKADIACLTSNAGLFAW 368  
 Qy 238 --RRQKEIRIQELGELHKQVLDVDSITGMSGP-----DFDNGTLAVL 277  
 Db 369 MDRLHLADRNSFSEIEILMHIIIDRVKLVNPSGSPCTEPGFRCFANMDDTLHVA 428  
 Qy 278 RGRWRYLMRSREITLGRATKDQIDVLSLEGPW 313  
 Db 429 LGRIDQFVSKNKKIVERKASENDVYONKSAKGLKW 464

RESULT 11  
 ID H101 ARATH STANDARD; PRT; 911 AA.  
 AC P42730; Q9LE57; Q8W4F2;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Heat shock protein 101.  
 GN HSP101 OR ATIG74310 OR F1017.2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae;  
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Leaf;  
 RX MEDLINE=95170291; PubMed=7866032;  
 RA Schirmer E.C., Lindquist S., Vierling E.;

RT "An Arabidopsis heat shock protein complements a thermotolerance  
 RT defect in yeast."  
 RL Plant Cell 6:1899-1909(1994).

RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20226114; PubMed=10760305;  
 RA Hong S.-W., Vierling E.  
 RT "Mutants of Arabidopsis thaliana defective in the acquisition of  
 RT tolerance to high temperature stress."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4392-4397(2000).

RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federpiet N.A., Kaul S.,  
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,  
 RA Miltescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sano H., Salzman J., Schwartz J.R., Shim P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utecherack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana."  
 RL Nature 408:816-820(2000).

RP SEQUENCE OF 1460 FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,  
 RT "The Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
 RT SSP consortium (Salk/Stanford/PGSC)."  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THERMOTOLERANCE.  
 CC -1- INDUCTION: BY HEAT SHOCK.  
 CC -1- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC -----  
 DR EMBL: U13949; AAA67927.1; -  
 DR EMBL: AF18796; AAF6423.1; -  
 DR EMBL: AC020579; AAS52410.1; -  
 DR EMBL: AY062596; AAL32674.1; -  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR003959; AAA\_ATPase\_cent.  
 DR InterPro: IPR001270; Chaprinin\_clpA/B.  
 DR InterPro: IPR004176; Clp\_N.  
 DR Pfam: PF00004; AAA; 2.  
 DR Pfam: PF02861; Clp\_N; 2.  
 DR PRINTS: PR00300; CLP\_PTEASEA.  
 DR ProDom: PD000739; GSP1\_E; 1.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00870; CLPB\_1; 1.  
 DR PROSITE: PS00871; CLPB\_2; 1.  
 KM Chapterone; Heat shock; ATP-binding; Repeat.  
 FT DOMAIN 164 410  
 FT DOWAIN 532 723  
 FT NP\_BIND 207 214  
 FT NP\_BIND 606 613  
 ATP (POTENTIAL).  
 ATP (POTENTIAL).



```

CC FT CONFLICT 141 141 V -> F (IN REF. 4).
CC FT CONFLICT 595 595 P -> A (IN REF. 1).
CC SQ SEQUENCE 911 AA; 101294 MW; 191EC1853B0C4C69 CRC64;
CC Query Match 5.0%; Score 99.5; DB 1; Length 911;
CC Best Local Similarity 23.8%; Pred. No. 17;
CC Matches 63; Conservative 42; Mismatches 97; Indels 63; Gaps 11;
CC
CC QY 105 ALLYPPVSKLACQAMROLHPFAIAIOSKALFSKAE---EQLLSVSGSTQPTLETFPD 161
CC DB 37 ALISDPT-----GIFPQISAGENNAQSAERVINQALKKLPQSQSPP-----DD 82
CC QY 162 LHHRPDAFYLTAKALQAHQMLKQVYVLEDOFQVPLPKGQDVLFSDADLIDDSKL 221
CC DB 83 I-----PASSSLKIVRRQAQAQKSGDTHLAVDQIM-----GLEDSQI 123
CC QY 222 KDMRDEVLHEHLMVADROKREIRQLDEQLHKQVLYVDSITGMSDFDNOQLAVL---- 277
CC DB 124 RDLNLEV-----GVATARKSEVEKLRGKEGK---KVESASG---DTNFQALKTVGRDL 171
CC QY 278 ---RRRMVRYLMRSREI-----TLGRATDNOQIDVDLSLEGPAWKISRKQVYIKLKNMGD 329
CC DB 172 VEOAGKLDPVIGRDEIRRVVRLSRRTKNIPV---LIGEPGVKTAVVEGLAQRIVKGD 228
CC QY 330 PFIANGRRPYIDGRPVLCGSKMR 354
CC DB 229 VENSITDVRLSLDMGALVAGAKYR 253
CC
CC RESULT 12
CC T2D1 DROME STANDARD; PRT; 2068 AA.
CC ID P51123;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Transcription initiation factor TFIIID 230 kDa subunit (TAFII-230)
CC DE (TAFII250) (TBP-associated factor 230 kDa) (P230).
CC GN TAF250.
CC OS Drosophila melanogaster (Fruit fly).
CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
CC OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
CC OC Muscomorpha; Bphidroides; Drosophilidae; Drosophila.
CC OX NCBI_TaxID=7227;
CC RN [1]
CC RP SEQUENCE FROM N.A., AND SEQUENCE OF 63-75 AND 540-546.
CC RX MEDLINE=93279463; PubMed=8504928;
CC RA Kokubo T., Gong D.-W., Yamashita S., Horikoshi M., Roeder R.G.,
CC RA Nakatani Y.;
CC RT "Drosophila 230-kD TFIIID subunit, a functional homolog of the human
CC RT cell cycle gene product, negatively regulates DNA binding of the TATA
CC RT box-binding subunit of TFIIID."
CC RL Genes Dev. 7:1033-1046(1993).
CC CC -I- FUNCTION: MAY PLAY AN ESSENTIAL ROLE IN TFIIID ASSEMBLY BY
CC CC INTERACTING WITH BOTH TBP AND OTHER TAF, AS WELL AS SERVING TO
CC CC LINK THE CONTROL OF TRANSCRIPTION TO THE CELL CYCLE. ESSENTIAL FOR
CC CC PROGRESSION OF THE G1 PHASE OF THE CELL CYCLE. POSSESSES DNA-
CC CC BINDING ACTIVITY. IS A NEGATIVE REGULATOR OF THE TATA BOX-BINDING
CC CC ACTIVITY OF TBP.
CC CC -I- SUBUNIT: TF2D IS COMPOSED OF TBP AND A VARIETY OF TBP-ASSOCIATED
CC CC FACTORS.
CC CC -I- SUBCELLULAR LOCATION: Nuclear.
CC CC -I- SIMILARITY: CONTAINS 2 BROMODOMAINS.
CC CC -I- SIMILARITY: CONTAINS 1 HMG BOX.
CC CC -I- SIMILARITY: TO HUMAN TAFII-250 (CCG1). SOME TO S.POMBE TAFII-111
CC CC AND TO S.CEREVISIAE TAF145.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC CC or send an email to license@ieb-sib.ch).
CC CC -----
CC CC EMBL; S61883; AAB26991.2; -.
CC CC DB HSBP; Q92831; 1B91.
CC CC DR TRANSFAC; T02119; -.
CC CC DR FLYBASE; FBGN0010355; TAF250.
CC CC DR InterPro; IPR000637; AT hook.
CC CC DR InterPro; IPR001487; Bromodomain.
CC CC DR InterPro; IPR001878; ZnF CCHC.
CC CC DR Pfam; PF00439; Bromodomain; 2.
CC CC DR Pfam; PF02178; AT hook; 1.
CC CC DR PRINTS; PR00503; BROMODOMAIN.
CC CC DR SMART; SM00384; AT hook; 1.
CC CC DR SMART; SM00297; BROMO; 2.
CC CC DR SMART; SM00343; ZnF C2HC; 1.
CC CC DR PROSITE; PS00633; BROMODOMAIN 1; 2.
CC CC DR PROSITE; PS50014; BROMODOMAIN 2; 2.
CC CC KW Bromodomain; Nuclear protein; DNA-binding; Cell cycle; Repeat;
CC CC KW Transcription regulation; Phosphorylation.
CC CC KM DNA BIND 1247 1360 HMG BOX (POTENTIAL).
CC CC FT DOMAIN 1445 1451 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC CC FT DOMAIN 1490 1560 BROMODOMAIN 1.
CC CC FT DOMAIN 1612 1682 BROMODOMAIN 2.
CC CC FT DOMAIN 1995 2068 GLN-RICH.
CC CC FT VARIANT 575 575 P -> S.
CC CC SQ SEQUENCE 2068 AA; 232494 MW; AD6A5ABF28B59531 CRC64;
CC
CC Query Match 5.0%; Score 99.5; DB 1; Length 2068;
CC Best Local Similarity 21.5%; Pred. No. 51;
CC Matches 68; Conservative 51; Mismatches 100; Indels 97; Gaps 15;
CC
CC QY 5 GGVEPGRCS-----GSEPSSEKKVSKAPSTPVPSPSP-----PAPGLTKGVKSK 51
CC DB 1000 GRADPTGCEGFSYVRVNPKPQYKEBQESQ-PKASVGTGTDALRLPLQRAKELLROK 1058
CC QY 52 QPLQYTKDLGRKWPANDLLINAVLQTDNLTSVHLGV-KFS--CRFTLREVERVYALLY 108
CC DB 1059 VVEEIKKLKLSRW---EVIDVRLTSTEKAKGEGMDKFSGNKFSIAHQERY----- 1109
CC QY 109 DPTVSKLACQAMROLHPFAIAIO---SKALFSAEQLSKVG----- 149
CC DB 1110 ---KEBQRFIDQNVLSSEVLSTDEASSSEBSDEELKLNLMNLSNKKSTQ 1164
CC QY 150 -STSQPTLETPODLL-----HRHPDAFYLTAKALQAHQML----- 186
CC DB 1165 LSRREBEILR-DELLRQIDEEHGSGSGAGAKGKDXDPGQMLATNNGRILATRTF 1223
CC QY 187 -----KOYVLEDOFQVPLPKGQDVLFSDADLIDDSKLKDMRDEVLHEHLMVADRRQ- 240
CC DB 1224 RNDQKEYTRVE--IVRQPVIDAVI-----KIRTKDEQFIKCPATLDEOOK 1269
CC QY 241 ---KEIRQLDEQLHK 253
CC DB 1270 EEMKREKRRIQELRLR 1285
CC
CC RESULT 13
CC PYR1 SCHPO STANDARD; PRT; 2244 AA.
CC ID Q09794;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE URAL protein (includes: Glutamine-dependent carbamoyl phosphate
CC DE synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2)).
CC GN URAL OR SPAC22G7.06C.
CC OS Schizosaccharomyces pombe (Fission yeast).
CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC OC Schizosaccharomycetes.
CC OX NCBI_TaxID=4896;
CC RN [1]
CC RP SEQUENCE OF 22-2244 FROM N.A.

```



RC STRAIN=972; PubMed=8590465;  
RX MEDLINE=96020160; PubMed=8590465;  
RA Lollier M., Jaquet L., Nedeva T., Lacroite F., Potier S.,  
RA Souciet J.-L.;  
RT "as in *Saccharomyces cerevisiae*, aspartate transcarbamoylase is  
RT assembled on a multifunctional protein including a dihydroorotase-like  
RT cryptic domain in *Schizosaccharomyces pombe*.";  
RL Curr. Genet. 28:138-149(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11893360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Scourse J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,  
RA Weljens I., Vanstreels E., Rieger M., Scheider M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
RA Bozsym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huret S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerretti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;  
RT "the genome sequence of *Schizosaccharomyces pombe*.";  
RL Nature 415:871-880(2002).  
RC -1- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING THREE  
CC ENZYMATIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPSASE,  
CC AND ATCASE) (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +  
CC phosphate + glutamate + carbamoyl phosphate + L-aspartate = phosphate  
CC + N-carbamoyl-L-aspartate.  
CC -1- PATHWAY: Pyrimidine biosynthesis; first step.  
CC -1- PATHWAY: Pyrimidine biosynthesis; second step.  
CC -1- MISCELLANEOUS: GATASE (GLUTAMINE AMIDOTRANSFERASE) AND CPSASE  
CC (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE  
CC GLUTAMINE-DEPENDENT CPSASE (GP-CPSASE) (EC 6.3.5.5).  
CC -1- MISCELLANEOUS: IN EUKARYOTES EC 6.3.5.5 IS SYNTHESIZED BY TWO  
CC PATHWAY-SPECIFIC (ARGININE AND PYRIMIDINE) UNDER SEPARATE CONTROL.  
CC -1- SIMILARITY: THE CPSASE DOMAIN IS SIMILAR TO OTHER CPSASES.  
CC -1- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE DHQASE FAMILY.  
CC DECEPTIVE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
CC EMBL, X81841; CAA57433.1; -  
DR EMBL, Z54328; CAA91130.1; -  
DR HSSP, P00968; IABX.  
DR InterPro, IPR002029; Asp/Orn Cotransf.  
DR InterPro, IPR002082; Asp carbinyltransf.  
DR InterPro, IPR001117; CPG\_GATase.  
DR InterPro, IPR000901; CPSase.  
DR InterPro, IPR002474; CPSase\_sm\_chain.

QY	288	SREITLGRATKDNQIDV	305
QY	15	SEPSSEKKYKSKAPSTVPPS-PAPAGGLTKRVKSKQ-----PQVTKDLGRMKPA	66
QY	67	NDLL-LINA-----VLQTNLTSVHLGVFSCFTLREVOERWYALLVYPVLSKLAC	117
QY	1923	HELVQILNSSFYRKHITSVQVTRSDHLVFLPALAHQRIIVERGVV--DLCTYKLLCT	1980
QY	118	-----QAMQLHPREALIAIQSLKSPKAE-----QLSKVGST--SOPT	155
QY	1981	MFFPSTRTSSSFDAQMRLGKVVAVATASASVVKGSLADTIRTLGCGDAIVLRHS	2040
QY	156	LETFDDLIH-----RHPDAFYL-----ATA	176
QY	2041	IESARIANFSVPVLIINGNGSKHEPFAFLDYLTIREELGSVNGLTITTFIGDLKYGRIV	2100
QY	177	KALQ-----AHWQLMKOYVLEDOQVPLPRKGOV-----LNFSDAEDLIDSKLDMRDEVL	229
QY	2101	HSLAHLAFMH-VELHLVSPEDLALPDDVKDIRANGANFLIEHRL-----TKEVYA	2151
QY	230	EHELMVADROKREIROLEQELHKQVILVDSITGMSSPDFNQTLAVLGR--MYRYLMR	287
QY	2152	QSDVLYCTRVQKERFASVDE-----YEKLKDSFI-----VNSVLASAKSHCTVHGLPLR	2201
QY	288	SREITLGRATKDNQIDV	305

Db 2202 NREIS-----EEVDFD 2212

RESULT 14

MUTS\_AOUPY STANDARD; PRT; 855 AA.

AC P70755;

DT 15-JUN-1998 (Rel. 36, Last Created)

DT 15-JUN-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE DNA mismatch repair protein mutS.

GN MUTS.

OS Aquifex pyrophilus.

OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;

CC Aquifex.

CC NCB1\_taxid=2714;

CC [1]

CC SEQUENCE FROM N.A.

CC Weimur J.G., Rosenfeld A., Wong D.M.;

CC "Hyperthermophilic Muts proteins: isolation, characterization and

CC enhancement of PCR specificity.";

CC Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.

CC -!- FUNCTION: This protein is involved in the repair of mismatches in

CC DNA. It is possible that it carries out the mismatch recognition

CC step. This protein has a weak ATPase activity (By similarity).

CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL; U71154; AAB16998.1; -

CC InterPro: IPR000432; Muts\_C.

CC InterPro: IPR002863; Muts\_N.

CC DR Pfam; PF00488; Muts\_C; 1.

CC DR Pfam; PF01624; Muts\_N; 1.

CC DR PRODOM; PD001263; Muts\_C; 1.

CC DR SMART; SM00534; Muts\_C; 1.

CC DR SMART; SM00533; Muts\_N; 1.

CC DR TIGRfam; TIGR01070; mutS1; 1.

CC DR PROSITE; PS00486; DNA\_MISMATCH\_REPAIR\_2; 1.

CC DR DNA repair; ATP-binding; DNA-binding.

CC NP BIND 607 614 ATP (POTENTIAL).

CC FT NP BIND 607 614 ATP (POTENTIAL).

CC SQ SEQUENCE 855 AA; 97656 MW; 5A8C8BEE234E8A CRC64;

Query Match 4.9%; Score 98.5; DB 1; Length 855;

Best Local Similarity 21.4%; Pred. No. 18;

Matches 83; Conservative 60; Mismatches 133; Indels 111; Gaps 18;

Oy 24 KVSCKPSTVPSPAP--APGLTKRVKSKOPQOVTKDL-GRWKPNADLLINAVLOTND 80

Db 236 KATQGSFTPLIKRPKYVDEGVKDLKAVKGLIEITSEGR---KDLSTLRKV---D 287

Oy 81 LTVSHLVGVKSCRFLL-----REVOERWYALLVDVTSIKLACQMRQLHPREALA 130

Db 288 RLVTGNG--RRRLRFLNPFRIERIRKQOE-----AVBELINK-----REVLANE 331

Oy 131 IQSKALFSAE-EQLISKVSGTSQPTLETFQDLIHRHPAFIARAKALQAHQWQMKRY 189

Db 332 IR-KTLEGMSDERLVRIS--NMASPRELIH-----LKNSLRAEELRKTL 376

Oy 190 YLEEDQVQPLPKGDQVLFNSDAEDLIDSKLQMDDEVLEHELMV-----ADRRQREIR 245

Db 377 SLIDSEIFKEIR--GSLNLNKAADLIDKTLVDDPPLHVKEGGLIRPGVNAVILDEIRFR 434

Oy 246 Q-LEOELHKVQLVDSITGMS-----PDF--DNQTLAVLNG 279

Db 435 EBAEKLIKVEKKKKETQISQIKGYNKVMGYIIEVTKANYKYVEHFRRRQTLSNAR 494

Oy 280 RMVRYIAMSREITLGRATKNOIDVLDL-----SLEGPAN 313

Db 495 YTTTELQRLERKILISAQRINLELEYRLREBEVVKELDKVGNNAVLIGEVDYIQSLAW 554

Oy 314 KISRQGVIKLKNNGDFPIANERRPI 340

Db 555 IALEKGWVKPEVHEGVEYLIEEGKHVP 581

RESULT 15

DYH9\_HUMAN STANDARD; PRT; 4486 AA.

AC Q9NYVC9; Q9NO28; O95494;

DT 16-OCT-2001 (Rel. 40, Last Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ciliary dynein heavy chain (Axonemal dynein heavy chain) (Dynein heavy chain 9).

GN DNAB9 OR DNAB17L OR DNAB1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CC NCB1\_taxid=9606;

CC [1]

CC SEQUENCE FROM N.A.

CC Reed W., Moats-Staats B.M., Carson J.L., Leigh M.W., Collier A.M.;

CC "A ciliary dynein heavy chain whose expression is upregulated in

CC differentiating airway epithelium.";

CC Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.

CC [2]

CC SEQUENCE FROM N.A.

CC TISSUE=Nasal epithelium;

CC Bartoloni L., Bloun J.L., Maiti A., Sainsbury A., Rossier C.,

CC Gehrig C., She J.X., Marron M.P., Lander E., Meeks M., Chung E.,

CC Jorissen M., Scott H.S., Delozier Blanchet C.D., Gardiner M.,

CC Antonarakis S.E.;

CC "Axonemal beta heavy chain dynein DNAB9: cDNA sequence, genomic

CC structure and investigation of its role in primary ciliary

CC dyskinesia.";

CC Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.

CC [3]

CC SEQUENCE OF 1874-1974 FROM N.A.

CC TISSUE=Nasal polyps;

CC Maiti A.K., Maitel M.G., Jorissen M., Volz A., Ziegler A.,

CC Bouvagnet P.;

CC "Chromosomal localization of human dynein heavy chain genes.";

CC Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.

CC [4]

CC CHARACTERIZATION.

CC MEDLINE=20558134; PubMed=1104725;

CC Reed W., Carson J.L., Moats-Staats B.M., Lucier T., Hu P.C.,

CC Brighton L., Gambling T.M., Huang C.H., Leigh M.W., Collier A.M.;

CC "Characterization of an axonemal dynein heavy chain expressed early in

CC airway epithelial ciliogenesis.";

CC Am. J. Respir. Cell Mol. Biol. 23:734-741(2000).

CC -!- FUNCTION: FORCE GENERATING PROTEIN OF RESPIRATORY CILIA. PRODUCES

CC FORCE TOWARDS THE MINDS ENDS OF MICROTUBULES. DYNEIN HAS ATPASE

CC ACTIVITY.

CC -!- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF

CC INTERMEDIATE AND LIGHT CHAINS.

CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL; AF257737; AAF69004.1; -

CC DR EMBL; AJ404468; CAB94756.1; -

CC DR EMBL; AJ132088; CAA10561.1; -

DR Genew; HGNC:2953; DNAB9.  
DR MIM; 603330; -.  
DR InterPro; IPR004273; Dynein heavy.  
DR Pfam; PF03028; Dynein heavy; 1.  
KM Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.  
FT DOMAIN 381 410 COILED COIL (POTENTIAL).  
FT DOMAIN 504 529 COILED COIL (POTENTIAL).  
FT DOMAIN 639 662 COILED COIL (POTENTIAL).  
FT DOMAIN 752 823 COILED COIL (POTENTIAL).  
FT DOMAIN 1326 1355 COILED COIL (POTENTIAL).  
FT DOMAIN 3051 3154 COILED COIL (POTENTIAL).  
FT DOMAIN 3285 3341 COILED COIL (POTENTIAL).  
FT DOMAIN 3640 3675 COILED COIL (POTENTIAL).  
FT NP\_BIND 1870 1877 ATP (POTENTIAL).  
FT NP\_BIND 2151 2158 ATP (POTENTIAL).  
FT NP\_BIND 2478 2485 ATP (POTENTIAL).  
FT NP\_BIND 2825 2832 ATP (POTENTIAL).  
FT CONFLICT 2505 2505 L -> V (IN REF. 2).  
FT CONFLICT 3678 3678 T -> A (IN REF. 2).  
FT CONFLICT 4374 4374 I -> M (IN REF. 2).  
SQ SEQUENCE 4486 AA; 511927 MW; 996EDFDEB0B3EB1 CRC64;

Query Match 4.9%; Score 98.5; DB 1; Length 4486;  
Best Local Similarity 21.3%; Pred. No. 1.8e+02;  
Matches 59; Conservative 47; Mismatches 116; Indels 55; Gaps 11;

QY 127 AIAIQSALFSKAEBSLSTKVGSTGPTLET-----FQDLLHRHPD-----AF 170  
Db 1372 AVAEIQNPRIERIRHQRLMQATGVSTMDQDTTAAHLQLQHLHYEDEVRGIVDKAKEM 1431  
QY 171 YIARTAKALQAHWQIMKQYLLLEDQTVQPLPKGQVLNFSDAEDLIDSKLDMRDEVLE 230  
Db 1432 GMEKTLKELQTTWAGMERQY-----EPHRTNVPILCSD-EDLI-----EYLE 1473  
QY 231 -HELMVADRQKREIRQLBQLHKQVLVDSITGMSPPDFNQTL-----AVLRGRWRY 284  
Db 1474 DNQVLOLWMSKYVAFLEEVSGWQKLSYDAVISIMFEVQRTWTHLESIFTG----- 1528  
QY 285 LMRGREITLGRATDNQID-VDSLSEGPWKISRKQGYIKLKNNGDFPIANEGRRPIYID 343  
Db 1529 --SEDIRAQLPQDSKRREGIDIDFKELAYDAQKIPNVAVQTTNKPGLYEKLED-----IQ 1580  
QY 344 GRPVLCGSKWRLSNNSVEIASLRFVFLINODLLALI 380  
Db 1581 GRCLIC-EKALAEYLDTRKLAFFRRYFLSSSDLLDIL 1616

Search completed: March 27, 2003, 05:20:21

Job time : 49 secs

**This Page Blank (uspto)**